DOCUMENT RESUME

ED 093 800 SP 008 120

AUTHOR Meublat, Guy
TITLE "DEP'ART." Un Modele de Prevision des Departs des Enseignants. Documents Demographie Scolaire 9-16 ("DEP'ART." A Model for Predicting Teacher Attrition. Scholastic Demographic Document 9-16).
INSTITUTION Quebec Dept. of Education, Quebec.
PUB DATE Apr 74
NOTE 109p.; Text only in French

EDRS PRICE MF-$0.75 HC-$5.40 PLUS POSTAGE
DESCRIPTORS *Dropout Identification; *Educational Research; *Faculty Mobility; *Teacher Persistence; Teacher Supply and Demand
IDENTIFIERS *Quebec

ABSTRACT

This document forms part of a research project initiated by the Ministry of Education in Quebec and designed to forecast teacher demand over the next 15 years. It analyzes the problem of identifying potential teacher dropouts by means of a statistical model which provides simulations of various hypotheses and which can be easily revised by the incorporation of new data. The observation period covered the four academic years 1967/68, 1968/69, and 1970/71. Chapter one provides the framework on which the model was constructed and includes statistics for three major categories: public and private sectors; nursery, elementary, secondary, and college levels; and sex. Other variables considered include professional training, matrimonial status, age, experience, religious affiliation of school, and type of region (urban or rural). Chapter two describes the development of the model, using age as the significant variable. Chapter three sets out the results obtained from the model and their application to future plans. Data used in the research project are included in the document. (MBM)

ED 093800

GOUVERNEMENT DU QUÉBEC
MINISTÈRE DE L'ÉDUCATION
DIRECTION GÉNÉRALE DE LA PLANIFICATION

DOCUMENTS

DÉMOGRAPHIE SCOLAIRE **9 - 16**

" DÉP'ART"

un modèle de prévision des départs d'enseignants

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SP008 120

AVRIL 1974

GUY MEUBLAT

D. 1183

28-1151

Je tiens à remercier MM. Bernard GENET et Claude-Guy SURPRENANT qui m'ont apporté, l'un dans le domaine de l'informatique, l'autre dans celui de la statistique, une collaboration appréciée.

G.M.

TABLE DES MATIERES

	Page
INTRODUCTION.....	2
CHAPITRE 1. LES DEPARTS D'ENSEIGNANTS: MESURES DU PHENOMENE.....	7
1. Evaluation globale.....	7
2. Influence de l'âge et de l'expérience.....	14
2.1 Le rôle de l'âge.....	14
2.2 Le rôle de l'expérience.....	21
3. Quantification de l'influence de certaines variables qualitatives.....	27
3.1 Méthodologie.....	27
3.2 L'application.....	29
3.3 Les résultats.....	30
3.4 Interprétation.....	37
3.4.1 Les équations secondaires.....	37
3.4.2 L'équation principale.....	38
CHAPITRE 2. LE MODELE "DEP'ART".....	41
1. Estimation des probabilités de départ selon l'âge.....	42
1.1 La technique de régression.....	42
1.2 Les résultats.....	46
2. Approximation de la structure par âge.....	52
2.1 Fonction de Pearson type "transition no 3" (ou "Pearson type 3").....	56
2.2 Fonction de Pearson "type 1".....	58
2.3 Les résultats.....	61

Table des matières
(suite)

Page

CHAPITRE 3. RESULTAT DES SIMULATIONS..... 69

1. Informatisation du système de calcul..... 69

2. Initialisation des variables nécessaires à la prévision..... 69

2.1 Les variables "Structurelles"..... 71
2.2 Les variables touchant l'âge..... 79

3. Les résultats..... 83

APPENDICE STATISTIQUE SOMMAIRE..... 89

ANNEXE..... 94

LISTE DES TABLEAUX

Page

TABLEAU 1. Nombre de départs parmi les enseignants; Secteur Public (entre parenthèses, l'effectif de base)..... 8

TABLEAU 2. Nombre de départs parmi les enseignants; Secteur Privé (entre parenthèses, l'effectif de base)..... 9

TABLEAU 3. Taux brut de départ (Secteur public)..... 11

TABLEAU 4. Taux brut de départ (Secteur privé) 12

TABLEAU 5. Coefficient de corrélation entre l'âge et l'expérience..... 15

TABLEAU 6. Age moyen des départs - Secteur public (entre parenthèses: l'âge moyen de l'effectif de base)..... 16

TABLEAU 7. Age moyen des départs - Secteur privé (entre parenthèses: l'âge moyen de l'effectif de base)..... 17

TABLEAU 8. Variance de l'âge des départs - Secteur public (entre parenthèses: la variance de l'effectif de base)..... 18

TABLEAU 9. Variance de l'âge des départs - Secteur privé (entre parenthèses: la variance de l'effectif de base)..... 19

Liste des tableaux
(suite)

Page

TABLEAU 10. Expérience moyenne des départs - Secteur public (entre parenthèses: l'expérience moyenne de l'effectif de base)..... 22

TABLEAU 11. Expérience moyenne des départs - Secteur privé (entre parenthèses: l'expérience moyenne de l'effectif de base)..... 23

TABLEAU 12. Variance de l'expérience des départs - Secteur public (entre parenthèses: variance de l'effectif de base)..... 24

TABLEAU 13. Variance de l'expérience des départs - Secteur privé (entre parenthèses: variance de l'effectif de base)..... 25

TABLEAU 14. Equations des probabilités de départ - Simulation no. 1..... 49

TABLEAU 15. Equations des probabilités de départ - Simulation no. 2..... 50

TABLEAU 16. Equations des probabilités de départ - Simulation no. 3..... 51

TABLEAU 17. Probabilité de départ à certains âges selon les différentes simulations (Secteur public)..... 53

TABLEAU 18. Probabilité de départ à certains âges selon les différentes simulations (Secteur privé)..... 54

Liste des tableaux
(suite)

Page

TABLEAU 19. Comparaison des taux bruts de départ obtenus avec les fonctions de Pearson (Probabilité de départ: Simulation no. 1) De Fin 1967-68 à fin 1970-71. Secteur public..... 66

TABLEAU 20. Comparaison des taux bruts de départ obtenus avec les fonctions de Pearson (Probabilité de départ: Simulation no. 2) De Fin 1967-68 à fin 1970-71. Secteur privé..... 67

TABLEAU 21. Hypothèse d'évolution des clientèles scolaires..... 72

TABLEAU 22. Evolution des effectifs du personnel enseignant - Sexe masculin..... 77

TABLEAU 23. Evolution des effectifs du personnel enseignant - Sexe féminin..... 78

TABLEAU 24. Comparaison des taux bruts de départ obtenus à partir des 3 simulations - Secteur public..... 80

TABLEAU 25. Comparaison des taux bruts de départ obtenus à partir des 3 simulations - Secteur privé..... 81

TABLEAU 26. Comparaison du nombre de départs obtenu à partir des 3 simulations - Secteur public. 84

Liste des tableaux
(suite)

Page

TABLEAU 27. Comparaison du nombre de départs obtenu à partir des 3 simulations - Secteur privé..... 85

TABLEAU 28. Comparaison du nombre de départs obtenu à partir des 3 simulations - Secteur public & privé..... 86

LISTE DES GRAPHIQUES

Page

GRAPHIQUE 1. Evolution des taux bruts de départ.... 13

GRAPHIQUE 2. Comparaison des structures par âge des départs et de l'effectif de base..... 20

GRAPHIQUE 3. Comparaison des structures selon l'expérience entre les départs et l'effectif de base..... 26

GRAPHIQUE 4. Variables influençant le départ Secteur: Public Niveau: Mat. Sexe: Féminin..... 33

GRAPHIQUE 5. Variables influençant le départ Secteur: Public Niveau: Elém. Sexe: Masculin..... 34

GRAPHIQUE 6. Variables influençant le départ Secteur: Public Niveau: Elém. Sexe: Féminin..... 34

GRAPHIQUE 7. Variables influençant le départ Secteur: Public Niveau: Sec. Sexe: Masculin..... 35

GRAPHIQUE 8. Variables influençant le départ Secteur: Public Niveau: Sec. Sexe: Féminin..... 35

GRAPHIQUE 9. Variables influençant le départ Secteur: Public Niveau: Collégial Sexe: Masculin..... 36

Liste des graphiques
(suite)

Page

GRAPHIQUE 10. Variables influençant le départ Secteur: Public Niveau: Collégial Sexe: Féminin..... 36

GRAPHIQUE 11. Comparaison des probabilités de départ Secteur public, secondaire, masculin.... 48

GRAPHIQUE 12. Comparaison des répartitions par âge, observées et simulées, Public, secondaire, masculin, 1970-71..... 64

GRAPHIQUE 13. Ratio "Elèves/Enseignant" Secteur public 1967-68 / 1977-78..... 74

GRAPHIQUE 14. Ratio "Elèves/Enseignant" Secteur privé 1967-68 / 1977-78..... 74

GRAPHIQUE 15. Taux de féminité du personnel Secteur: public, 1967-68 / 1977-78..... 75

GRAPHIQUE 16. Taux de féminité du personnel Secteur: privé, 1967-68 / 1977-78..... 76

INTRODUCTION

Depuis deux ans environ, la Direction Générale de la Planification a entrepris de raffiner les prévisions de personnel, en analysant les différents mouvements qui provoquent l'évolution d'une population de ce type.
En ce qui concerne les enseignants proprement dits (à l'exclusion, donc, des membres de la direction, des cadres ou du personnel administratif), les mouvements se composent:

- des départs (enseignants quittant le système d'éducation)
- des arrivées (nouveaux enseignants sans expérience professionnelle)
- des retours (nouveaux enseignants ayant une expérience antérieure)
- des "migrants" (enseignants demeurant dans le système d'éducation mais changeant de fonction, ou de secteur, ou de niveau, ou de région, ou bien combinant plusieurs de ces changements)
- des "stables" (enseignants ne faisant l'objet d'aucun changement) qui, quoique le nom puisse en laisser penser, représentent bien un type de mouvement.

Parallèlement à une analyse chronologique détaillée de tous ces mouvements (1), la prévision de l'un de ces mouvements

(1) Cette étude est actuellement en cours à la D.G.P. et paraîtra vraisemblablement en 1974.

en particulier, à savoir les départs, a été entreprise. Les données nécessaires à une telle étude ont été obtenues du fichier dit de "mobilité du personnel", construit par le service de l'Informatique (S.I.M.E.Q.) en comparant deux à deux les rencensements annuels du personnel (2). Cette comparaison permet d'établir qu'un enseignant, présent la première année, ne l'est plus l'année suivante et de l'enregistrer, ainsi que toutes ses caractéristiques individuelles (sexe, âge, état matrimonial, état de vie (3), formation, secteur d'enseignement, niveau d'enseignement, confessionnalité de l'école, région, expérience dans l'enseignement) dans la catégorie des "départs". Cette méthode permet aussi, bien sûr, de comparer les caractéristiques de ce type de mouvement avec celles du personnel enseignant dans son ensemble.

Cette procédure, toutefois, présente une faille: elle ne permet pas de différencier, à l'intérieur des "départs", certaines causes comme le décès, la retraite, le changement de profession, le retour à l'inactivité ou le congédiement. Une étude précédente (4) a montré qu'en fait la mortalité était négligeable (parce que le personnel se concentre à des âges où le risque de décès est faible) et que les mises à la retraite, en plus d'être fortement dépendantes de l'âge, n'étaient

---

(2) Le plus ancien fichier utilisé est celui de 1967-68: la première période d'observation des mouvements vient donc de la comparaison des années 1967-68 et 1968-69; la dernière période disponible actuellement est 1970-71/1971-72.

(3) laïc ou religieux.

(4) cf. Guy Meublat: "Estimation des départs à la retraite des enseignants..." D.G.P., M.E.Q., Documents 9-10, août 1973.

pas très nombreuses. Enfin, argument non négligeable, c'est la seule procédure utilisable pour l'instant...

Cette prévision des départs, limitée aux enseignants, a fait l'objet d'un modèle statistique, finement baptisé "Dép'art" par les soins de l'auteur, dont la technique d'élaboration suit les règles traditionnelles de l'économétrie: établissement d'équations par la méthode dite de "régression", qui consiste à estimer les paramètres de fonctions simples en les confrontant avec des données empiriques. Le lecteur pourra consulter, à cet effet, l'appendice mathématique qui se trouve à la fin du document.

L'avantage d'un tel modèle, une fois que la phase technique d'élaboration est terminée (1), est de permettre une informatisation facile du procédé. Cette informatisation présente deux avantages importants, qui se traduisent par un gain de temps appréciable:

- elle permet la simulation de différentes hypothèses (nous en verrons une illustration dans le chapitre 3)
- elle facilite la révision des prévisions antérieures, en intégrant très rapidement toute nouvelle information relative à son objet (l'adjonction d'une nouvelle année de référence, par exemple).

---

(1) Cf. D.B. Suits: "... the compilation of an econometric model requires a certain degree of technical specialisation but, once constructed, any competent economist can apply it to policy analysis and economic forecasting" (in Forecasting and [illegible] econometric model, A.E.R., mars 1962)

Cette dernière qualité est essentielle, car elle doit permettre de pallier les faiblesses actuelles d'une telle prévision; faiblesses qui tiennent à :

- une période d'observation assez courte, qui va de la fin de l'année académique 1967-68 à celle de 1970-71 (soit 4 périodes) (1).

- la possibilité de certaines lacunes dans le recensement du personnel, notamment dans les années les plus anciennes; lacunes qui peuvent avoir pour conséquence de gonfler le nombre des départs, soit en ne retraçant pas correctement les enseignants qui migrent à l'intérieur du système (et qui sont alors considérés comme "départs" à leur institution d'origine), soit en omettant certaines institutions dans le recensement (2) (dont le personnel est alors compté dans les "départs" d'une année, et dans les "retours" deux ans plus tard).
Ces lacunes sont, en fait, marginales et ont été réduites par l'abandon des données antérieures à 1967-68 (2).

- la récence de certaines transformations de structure (création des C.E.G.E.P. passage de la 7ème année au secondaire, etc...) qui a pu influencer les comportements.

---

(1) défaut atténué par le fait que le modèle de prévision est avant tout un modèle à court terme.

(2) ceci est surtout valable pour le secteur privé et parfois pour le collégial car "l'obligation" de réponse y est moins forte.

On voit donc que la prévisions des départs d'enseignants (1), qui est donnée dans ce document (au chapitre 3), n'est que l'ébauche —temporaire— d'un travail de longue haleine; elle devra être jugée comme telle.

(1) la définition des enseignants est identique à celle qui est utilisée dans les brochures officielles depuis 1969-70.

Chapitre 1

LES DEPARTS D'ENSEIGNANTS: MESURES DU PHENOMÈNE

Trois catégories nous serviront de référence tout au long de notre étude : le secteur d'enseignement (public ou privé), le niveau d'enseignement (maternelle, élémentaire, secondaire, collégial) et le sexe (masculin, féminin). Il est évident, en effet, que des différences appréciables de comportement peuvent apparaître selon la combinaison de ces trois éléments.

Dans un premier temps, nous allons utiliser les compilations statistiques, tirées de notre fichier, pour connaître l'ampleur du phénomène, ainsi que ses caractéristiques vis-à-vis de l'âge et de l'expérience des enseignants. Ensuite, nous utiliserons une analyse de régression d'un type particulier (utilisant des variables dites "dichotomiques") pour essayer de mettre en évidence l'influence de certaines variables qualitatives (formation professionnelle, état matrimonial, statut, confessionnalité de l'école, type de région): nous en présenterons la méthodologie à ce moment.

1. Evaluation globale

Les tableaux 1 et 2 nous montrent que le nombre de départs, pour l'ensemble du système d'éducation, est très important: 13,140 à la fin de 1967-68 (18% de l'effectif total), 16,990 en 1968-69 (22% de l'effectif), 14,300 en 1969-70 (18.5% de l'effectif), et 10,360 en 1970-71 (13.5% de l'effectif).

TABLEAU 1: Nombre de départs parmi les enseignants;

Secteur Public (entre parenthèses, l'effectif de base)

Catégories / Années		1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	2 (11)	4 (17)	3 (15)	0 (2)
	Elémentaire	440 (3,424)	576 (3,635)	572 (4,013)	441 (3,813)
	Secondaire	2,145 (12,882)	3,349 (15,653)	3,076 (17,665)	2,305 (18,605)
	Collégial	261 (790)	641 (2,213)	1,000 (2,288)	366 (2,131)
SEXE FEMININ	Maternelle	256 (1,837)	454 (2,494)	392 (2,675)	356 (2,781)
	Elémentaire	4,617 (32,371)	5,471 (31,849)	4,302 (31,811)	3,443 (30,908)
	Secondaire	2,575 (12,141)	3,468 (13,034)	2,871 (13,460)	2,204 (14,222)
	Collégial	92 (238)	220 (574)	324 (755)	234 (891)

TABLEAU 2: Nombre de départs parmi les enseignants;

Secteur privé (entre parenthèses, l'effectif de base)

Catégories		Années 1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	2 (4)	2 (5)	3 (3)	0 (0)
	Elémentaire	51 (193)	78 (198)	69 (191)	48 (174)
	Secondaire	500 (2,091)	431 (1,239)	456 (1,282)	217 (931)
	Collégial	787 (2,640)	755 (1,715)	189 (490)	92 (437)
SEXE FEMININ	Maternelle	68 (196)	86 (196)	72 (173)	53 (147)
	Elémentaire	337 (1,143)	387 (1,132)	330 (958)	207 (874)
	Secondaire	432 (1,350)	525 (1,263)	442 (1,186)	239 (949)
	Collégial	574 (1,429)	544 (1,174)	276 (505)	155 (351)

Les tableaux 3 et 4 nous permettent de différencier ces chiffres: le taux brut de départ y est obtenu en divisant le nombre de départs dans chacune des catégories par l'effectif de base de ces mêmes catégories. Nous voyons que:

- La fréquence des départs est plus élevée au secteur privé qu'au secteur public.
- A l'intérieur des secteurs, la fréquence des départs est plus importante au secondaire et au collégial qu'à l'élémentaire.
- Les hommes partent généralement plus que les femmes (à l'exception du collégial, et du secondaire-privé)

Les graphique 1, représente l'évolution chronologique de tous les taux bruts. Toutefois, il faut prendre garde à son interprétation: le taux brut dépend non seulement de la fréquence des départs, mais aussi de la structure de l'effectif (par rapport à l'âge ou à toute autre variable de référence). Son évolution pourrait, dès lors, ne représenter qu'un changement dans cette structure (1). Il y a, en fait, très peu de chances que cela se produise dans les catégories où l'effectif est suffisament nombreux: on peut ainsi déduire de l'analyse de ce graphique que le mouvement de départ a atteint un maximum dans les années centrales (fin 1968-69 ou fin 1969-70)

---

(1) le cas rappelle, bien entendu, l'analyse de la fécondité avec la différence entre le taux brut de natalité et les taux de fécondité par âge ou groupe d'âges.

TABLEAU 3: Taux brut de départ

(Secteur public)

Catégories		Années 1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	.182	.235	.200	Effectif insuffisant
	Elémentaire	.129	.158	.143	.116
	Secondaire	.167	.214	.174	.124
	Collégial	.330	.290	.437	.172
SEXE FEMININ	Maternelle	.139	.182	.147	.128
	Elémentaire	.143	.172	.135	.111
	Secondaire	.212	.266	.213	.155
	Collégial	.387	.383	.429	.263

TABLEAU 4: Taux brut de départ

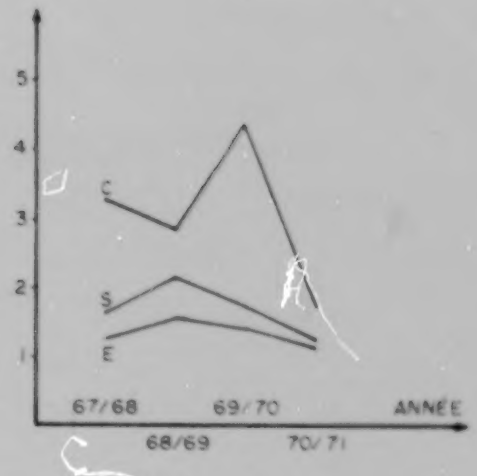
(Secteur privé)

Catégories		Années 1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	Effectif insuffisant	---	---	---
	Elémentaire	.264	.394	.361	.276
	Secondaire	.239	.348	.356	.233
	Collégial	.298	.440	.386	.211
SEXE FEMININ	Maternelle	.347	.439	.416	.361
	Elémentaire	.295	.342	.344	.361
	Secondaire	.320	.416	.373	.252
	Collégial	.402	.463	.547	.442

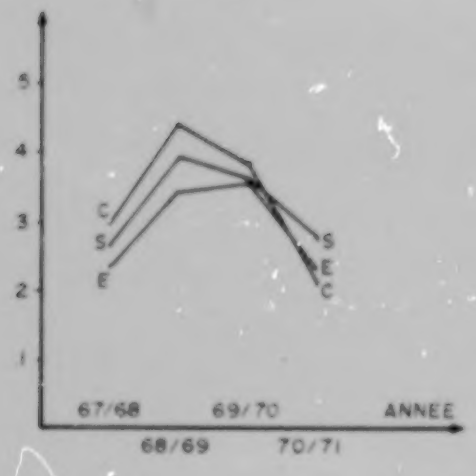
GRAPHIQUE I : ÉVOLUTION DES TAUX BRUTS DE DÉPART

M MATERNELLE E ELEMENTAIRE
S SECONDAIRE C COLLEGIAL

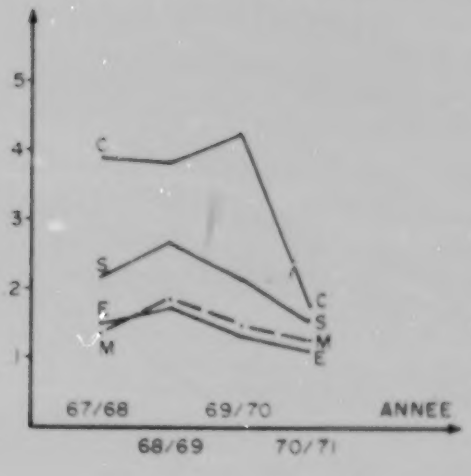
SECTEUR PUBLIC - SEXE MASCULIN



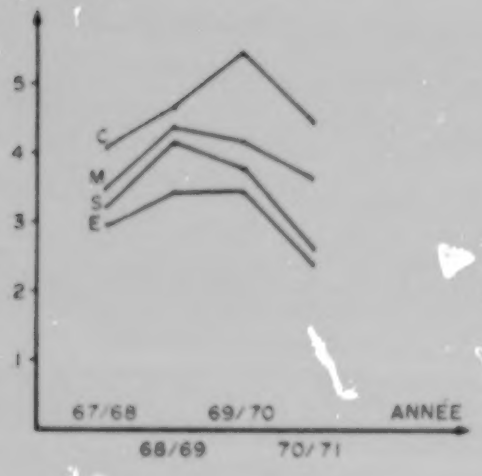
SECTEUR PRIVÉ - SEXE MASCULIN



SECTEUR PUBLIC - SEXE FÉMININ



SECTEUR PRIVE - SEXE FÉMININ



et son minimum, presque toujours, à la fin de 1970-71. Il est possible, donc, que la tendance soit au déclin du phénomène, mais il est impossible, étant donné le nombre restreint d'années d'observation, de l'affirmer.

2. Influence de l'âge et de l'expérience

L'idée de grouper ces 2 variables quantitatives n'est pas gratuite puisque, comme le montre le tableau 5, leur coefficient de corrélation linéaire est élevé. Elles auront donc une influence identique, ou presque, sur le phénomène des départs.

2.1 Le rôle de l'âge (1) (sans contrôle des autres variables)

Les tableaux 6 et 7 révèlent qu'il n'y a pas de très grandes différences entre le groupe des "Départs" et l'effectif de base (qui inclut d'ailleurs les départs), si ce n'est que les départs ont souvent une variance plus élevée (tableaux 8 et 9). Ce qui laisse penser que la distribution selon l'âge y est plus élevée aux extrémités de l'intervalle de définition.

Une telle analyse est confirmée par le graphique 2; graphique qui nous laisse prévoir également, que la fréquence

---

(1) Notre âge est repéré au 30 juin de l'année où peut se produire le départ (1968 pour 1967-68) alors que, dans les publications officielles, il est repéré un an plus tôt (1967 pour 1967-68); Si on veut retrouver les chiffres des publications, il suffira —puisque la moyenne est un opérateur linéaire— de retirer une année à notre âge moyen. La variance, elle, ne varie pas.

TABLEAU 5: Coefficient de corrélation entre l'âge et l'expérience.

Sexe / Secteur \ Niveau		Maternelle	Elémentaire	Secondaire	Collégial
SEXE MASCULIN	Public	-----	.84	.75	.69
	Privé	-----	.76	.74	.73
SEXE FEMININ	Public	.71	.79	.78	.75
	Privé	.58	.73	.73	.76

{Nota: Ces coefficients ont été calculés sur un nombre très important de "paires", allant de 680 (Privé-maternelle-féminin) à 126.[illegible]50 (Public-élémentaire-féminin)}

TABLEAU 6: Age moyen des départs - Secteur public

(entre parenthèses: l'âge moyen de l'effectif de base)

Catégories / Années		1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	25.0 (31.2)	27.0 (27.6)	25.7 (25.3)	---
	Elémentaire	31.6 (29.9)	30.4 (29.9)	29.4 (29.8)	30.3 (29.8)
	Secondaire	31.4 (34.5)	31.1 (31.4)	31.6 (31.5)	32.2 (31.9)
	Collégial	32.8 (34.3)	32.7 (33.3)	33.9 (33.5)	32.9 (32.7)
SEXE FEMININ	Maternelle	28.1 (27.8)	26.9 (27.6)	27.3 (27.7)	27.9 (28.3)
	Elémentaire	31.8 (31.4)	31.2 (31.7)	31.5 (32.2)	33.2 (32.8)
	Secondaire	32.3 (32.3)	31.9 (32.1)	32.1 (32.3)	31.3 (32.5)
	Collégial	34.9 (35.0)	33.6 (33.5)	34.2 (34.5)	32.1 (32.7)

TABLEAU 7: Age moyen des départs - Secteur privé

(entre parenthèses: l'âge moyen de l'effectif de base)

Catégories \ Années		1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	----	----	----	----
	Elémentaire	32.0 (34.5)	34.9 (34.1)	33.3 (32.3)	31.0 (31.8)
	Secondaire	34.8 (34.3)	37.7 (36.4)	37.4 (36.7)	35.9 (36.5)
	Collégial	35.9 (35.3)	35.8 (35.6)	36.3 (35.2)	35.8 (36.7)
SEXE FEMININ	Maternelle	34.1 (35.7)	35.0 (35.9)	32.1 (35.6)	31.3 (34.0)
	Elémentaire	34.9 (35.9)	35.9 (35.6)	34.5 (35.1)	31.6 (34.5)
	Secondaire	36.6 (38.1)	38.2 (38.4)	37.2 (37.9)	38.0 (37.6)
	Collégial	35.7 (36.2)	35.5 (36.2)	31.6 (32.7)	32.1 (34.2)

TABLEAU 8: Variance de l'âge des départs - Secteur public

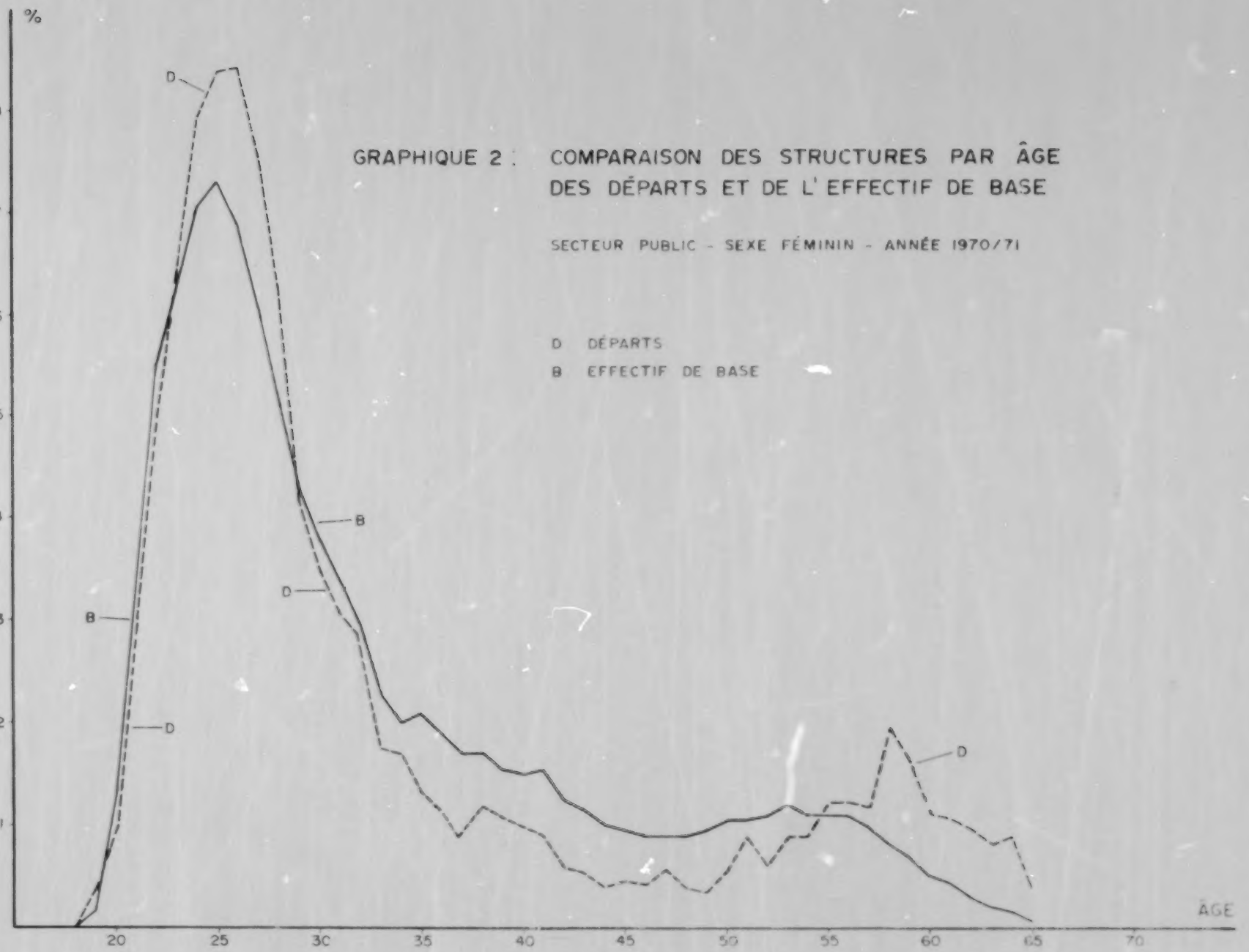
(entre parenthèses: la variance de l'effectif de base)

Catégories \ Années		1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	----	----	----	----
	Elémentaire	138.2 (92.9)	109.3 (87.8)	92.4 (83.4)	101.6 (79.9)
	Secondaire	94.8 (80.4)	88.1 (77.9)	97.0 (78.2)	102.8 (77.4)
	Collégial	105.0 (83.8)	87.0 (91.3)	89.5 (88.0)	90.6 (73.5)
SEXE FEMININ	Maternelle	84.2 (83.3)	57.5 (72.1)	66.8 (68.7)	77.4 (67.1)
	Elémentaire	158.5 (133.0)	133.0 (126.3)	139.8 (122.8)	158.8 (18.1)
	Secondaire	127.1 (116.9)	113.5 (110.7)	122.9 (109.9)	109.0 (106.1)
	Collégial	131.2 (135.6)	124.1 (118.6)	115.4 (110.6)	99.4 (98.1)

TABLEAU 9: Variance de l'âge des départs - Secteur privé

(entre parenthèses; la variance de l'effectif de base)

Catégories \ Années		1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	----	----	----	----
	Elémentaire	123.6 (162.2)	157.9 (149.0)	144.8 (126.2)	108.1 (107.8)
	Secondaire	147.6 (121.0)	173.8 (145.4)	158.5 (138.2)	159.7 (139.8)
	Collégial	108.4 (98.4)	102.0 (97.4)	112.1 (97.9)	123.9 (98.5)
SEXE FEMININ	Maternelle	142.4 (140.4)	145.4 (140.9)	134.4 (148.2)	126.0 (146.9)
	Elémentaire	165.0 (151.1)	169.3 (155.2)	149.5 (145.2)	118.9 (134.2)
	Secondaire	167.9 (166.5)	188.8 (172.7)	193.3 (170.4)	181.9 (166.4)
	Collégial	120.5 (121.5)	130.3 (126.6)	80.8 (94.8)	87.6 (100.8)



GRAPHIQUE 2 : COMPARAISON DES STRUCTURES PAR ÂGE DES DÉPARTS ET DE L'EFFECTIF DE BASE

SECTEUR PUBLIC - SEXE FÉMININ - ANNÉE 1970/71

D DÉPARTS

B EFFECTIF DE BASE

du départ sera plus élevée aux âges jeunes et aux âges élevés:

$$Y(x) = \frac{D(x)}{E(x)} = \frac{D \times U(x)}{E \times V(x)} = \frac{D}{E} \times \frac{U(x)}{V(x)}$$

[avec: Y(x)= fréquence du départ à l'âge x , D= nombre total de départs , E= effectif de base , U(x)= proportion des départs ayant l'âge x , V(x)= proportion de l'effectif de base ayant l'âge x]

Or, $\frac{D}{E}$ est une donnée ; donc, si $U(x) > V(x)$, $y(x) > \frac{D}{E}$ et réciproquement.

2.2 Le rôle de l'expérience (sans contrôle des autres variables)

Les tableaux 10,11,12,13, ainsi que le graphique 3, confirment l'analyse précédente — comme il fallait s'y attendre. On remarque simplement que, si la différence de structure est importante quand l'expérience est faible, on ne retrouve pas clairement, comme avec l'âge, cette différence lorsque l'expérience est forte. Ce qui semblerait indiquer, qu'à un âge donné, l'expérience peut, si elle dépasse un certain nombre d'années, constituer un frein au départ.

TABLEAU 10: Expérience moyenne des départs - Secteur public

(entre parenthèses: l'expérience moyenne de l'effectif de base)

Catégories / Années		1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	1.0 (4.4)	3.0 (4.9)	4.3 (3.5)	----
	Elémentaire	8.2 (7.1)	7.7 (7.2)	6.3 (6.9)	6.7 (6.7)
	Secondaire	7.7 (8.0)	6.9 (7.5)	6.3 (7.1)	7.0 (7.2)
	Collégial	6.8 (7.1)	6.4 (7.1)	5.1 (5.6)	5.5 (6.0)
SEXE FEMININ	Maternelle	5.7 (5.8)	5.2 (5.9)	5.7 (6.2)	5.6 (6.3)
	Elémentaire	9.8 (9.6)	9.4 (9.9)	9.1 (10.1)	10.0 (10.1)
	Secondaire	9.3 (9.7)	9.1 (9.6)	8.1 (9.2)	8.2 (9.1)
	Collégial	8.5 (10.0)	8.9 (8.9)	6.6 (7.7)	5.4 (6.3)

TABLEAU 11: Expérience moyenne des départs - Secteur privé

(entre parenthèses: l'expérience moyenne de l'effectif de base)

Catégories / Années		1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	----	----	----	----
	Elémentaire	8.2 (10.2)	10.0 (10.1)	6.1 (6.8)	7.0 (6.9)
	Secondaire	7.9 (7.8)	10.8 (10.4)	7.2 (9.0)	9.1 (10.4)
	Collégial	7.7 (8.0)	8.3 (8.5)	6.8 (7.5)	3.9 (8.6)
SEXE FEMININ	Maternelle	6.2 (8.6)	8.3 (9.1)	4.2 (6.4)	4.1 (6.2)
	Elémentaire	10.9 (12.0)	11.3 (11.5)	8.0 (8.8)	7.0 (8.8)
	Secondaire	11.3 (12.3)	11.4 (12.4)	8.3 (10.1)	10.7 (11.3)
	Collégial	10.0 (10.7)	9.6 (10.3)	4.3 (5.4)	3.9 (5.1)

TABLEAU 12: Variance de l'expérience des départs - Secteur public

(entre parenthèses: variance de l'effectif de base)

	Années / Catégories	1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	----	----	----	----
	Elémentaire	109.7 (70.0)	89.5 (66.8)	68.7 (64.7)	74.7 (58.0)
	Secondaire	80.0 (66.1)	65.5 (59.2)	61.1 (59.6)	70.2 (55.3)
	Collégial	63.4 (55.7)	59.5 (61.5)	55.5 (51.0)	56.1 (41.9)
SEXE FEMININ	Maternelle	29.8 (30.4)	20.0 (28.1)	38.1 (39.1)	32.3 (29.8)
	Elémentaire	87.7 (67.9)	72.3 (65.2)	82.2 (72.0)	88.2 (65.0)
	Secondaire	84.6 (79.3)	78.5 (75.6)	81.1 (78.7)	75.5 (72.2)
	Collégial	62.6 (91.1)	94.5 (89.4)	83.8 (89.0)	67.3 (68.3)

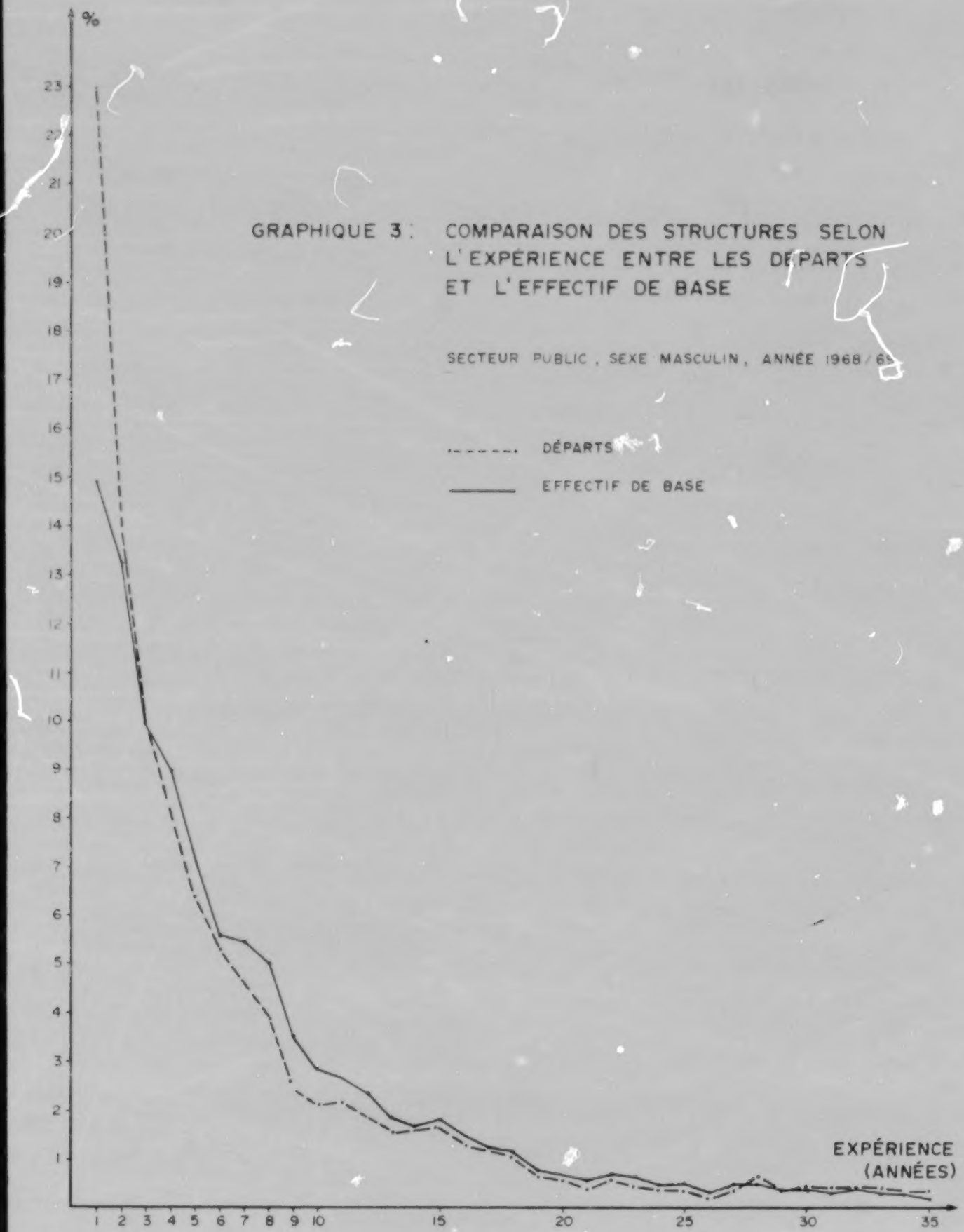
TABLEAU 13: Variance de l'expérience des départs- Secteur privé

(entre parenthèses: variance de l'effectif de base)

Catégories / Années		1967/68	1968/69	1969/70	1970/71
SEXE MASCULIN	Maternelle	----	----	----	----
	Elémentaire	114.7 (124.2)	101.7 (106.5)	87.1 (74.0)	76.1 (58.1)
	Secondaire	90.9 (83.7)	147.4 (121.1)	106.2 (109.7)	109.3 (114.8)
	Collégial	81.8 (70.7)	67.5 (66.8)	64.2 (64.2)	87.8 (83.7)
SEXE FEMININ	Maternelle	50.8 (67.7)	65.0 (60.5)	40.9 (56.7)	13.2 (43.4)
	Elémentaire	104.4 (102.0)	106.4 (100.2)	87.9 (84.9)	53.7 (77.2)
	Secondaire	138.0 (130.5)	144.0 (138.5)	135.7 (134.9)	142.6 (133.2)
	Collégial	103.7 (106.6)	104.3 (105.5)	35.1 (60.0)	27.9 (39.4)

GRAPHIQUE 3 : COMPARAISON DES STRUCTURES SELON L'EXPÉRIENCE ENTRE LES DÉPARTS ET L'EFFECTIF DE BASE

SECTEUR PUBLIC, SEXE MASCULIN, ANNÉE 1968/69



3. Quantification de l'influence de certaines variables qualitatives

3.1 Méthodologie

La mise en évidence du rôle d'une certaine variable qualitative, telle que l'état matrimonial, l'état de vie, la formation, etc..., peut être menée à partir de compilations statistiques traditionnelles. Mais cette démarche impose, pour éliminer les effets "parasites" provenant des autres variables, des recoupements extrêmement lourds. C'est pourquoi nous avons utilisé un modèle particulier de régression, à base de variables "dichotomiques" (1).

Est appelée "dichotomique", toute variable qui ne prend que 2 valeurs: 0 et 1, par exemple. Cette variable peut jouer le rôle de variable dépendante, ou de variable explicative; c'est à dire qu'une équation de régression pourra comporter une variable dépendante ("expliquée") dichotomique et

---

(1) -Pour plus de détails sur la technique utilisée, on pourra consulter: D.B. Suits: "Use of dummy variables in regression analysis", American Statistical Association Journal, no 52, dec. 1957, pp. 548-51.
-J. Johnston: "Econometric methods", chap. 6, pp. 176, Mc Graw-Hill, 2è ed., 1972.
-Goldberger: "Econometric theory", Wiley & Son, 1964.
-Draper & Smith: "Applied regression analysis", chap. 5 pp. 134 et 5 Wiley & Son, 1967.
-Un essai d'application au personnel enseignant a déjà été tenté par J.F. Gautrin: "Une évaluation quantitative des causes économiques de la mobilité géographique des enseignants", mémoire de maîtrise, département des sciences économiques. Université de Montréal, 1967.

des variables explicatives dichotomiques ou quantitatives:

$$\hat{Y} = a\,W + b\,X + c\,Z + E$$
$$(0,1)\quad (0,1)\quad (\text{quant.})$$

Dans ce cas, $\hat{Y}$ pourra être considéré comme la probabilité conditionnelle de réalisation d'un événement et pourra varier entre 0 et 1, même si les observations empiriques de Y — qui sont à la base de la régression— ne peuvent prendre que les valeurs 0 et 1.

Ce type de modèle exige, cependant, certaines précautions:

• la condition traditionnelle de l'homoscédasticité (constance de la variance de la distribution gaussienne des résidus) n'est pas respectée. Ne disposant pas des programmes informatiques permettant de corriger ce défaut (à l'aide d'une procédure semblable à celle que propose Goldberger), nous avons utilisé, telles qu'elles, les méthodes "classiques" de calcul. Ce qui est acceptable, vu l'utilisation que nous ferons de ce modèle.

• Il faut prendre garde, si on utilise les programmes conventionnels de calcul, de ne pas tomber dans ce que Johnston appelle "the dummy trap"(1), qui peut provoquer le non-sens de l'équation.

---

(1) Johnston, op.cit. Cette expression pourrait se traduire par "le piège de la variable dichotomique", p. 178.

• Les coefficients de corrélation produits par les programmes sont anormalement faibles et biaisés, parce qu'ils sont calculés à partir de la formule de Pearson. En effet les coefficients de corrélation impliquant des variables dichotomiques sont d'un type très particulier, appelé "bi-sérial" (entre une variable dichomotique et une variable quantitative), "tétrachorique" (entre 2 variables dichotomiques) ou coefficient "phi" (φ). Ne disposant pas des moyens informatiques pour corriger ce défaut, nous avons dû nous résoudre à ne pas utiliser les coefficients de corrélation dans l'analyse(1).

En utilisant le test de Snedecor (test de F), on sélectionne quand même des coefficients de régression non-biaisés et convergents.

3.2 L'application

Toutes les observations disponibles sur le fichier sont utilisées dans la régression, selon le secteur, le niveau et le sexe, avec les valeurs suivantes:

- Age
- Expérience
- Etat matrimonial { 0 = non célibataire. / 1 = célibataire.

(1) cf. L.T. Dayhaw:" Manuel de statistique", chap. 11 et 12, Université d'Ottawa, 1969.
J.P. Guilford & N.C. Perry:" Estimation of other coefficients of correlation from the phi coefficient", Psychometrika, vol.16, no 3, 1951.

- Formation 1 { 1 = Brevet A ou diplôme universitaire en pédagogie ou diplôme universitaire avec formation pédagogique.
  0 = Autre formation.
- Formation 2 { 1 = Brevet B ou équivalent.
  0 = Autre formation.
- Statut ou Etat de vie { 1 = Religieux.
  0 = Laïc.
- Région (1) { 0 = Montréal ou Québec (grandes agglomérations)
  1 = Autre région.
- Confessionnalité { 0 = école catholique.
  1 = école protestante.
- Taux de variation de la clientèle (qui dépend de l'année).
- Départ { 1 = oui.
  0 = non.

On voit qu'un enseignant, n'ayant aucune des formations citées, peut avoir la valeur 0 à la fois pour "Formation 1" et pour "Formation 2", ce qui évite la possibilité du "dummy trap".

3.3 Les résultats

Travaillant par secteur-niveau-sexe, et ayant exclu les "maternelle-masculin" pour insuffisance d'effectif, nous disposons donc de 14 équations de régression; la variable

(1) Il s'agit de la région administrative scolaire.

dépendante est la probabilité de départ et les variables explicatives sont dichotomiques ou quantitatives[1]. Ces équations permettent d'analyser l'influence d'une variable, toutes choses égales d'ailleurs: le signe du coefficient de régression indiquant dans quel sens varie la probabilité de départ lorsque, toutes les autres variables étant fixées, la variable dichotomique est égale à 1. Toutefois:

- L'équation donne autant de valeurs qu'il y a de combinaisons entre les variables.

- Les coefficients de régression dépendent de l'unité de mesure des variables, en particulier des variables quantitatives.

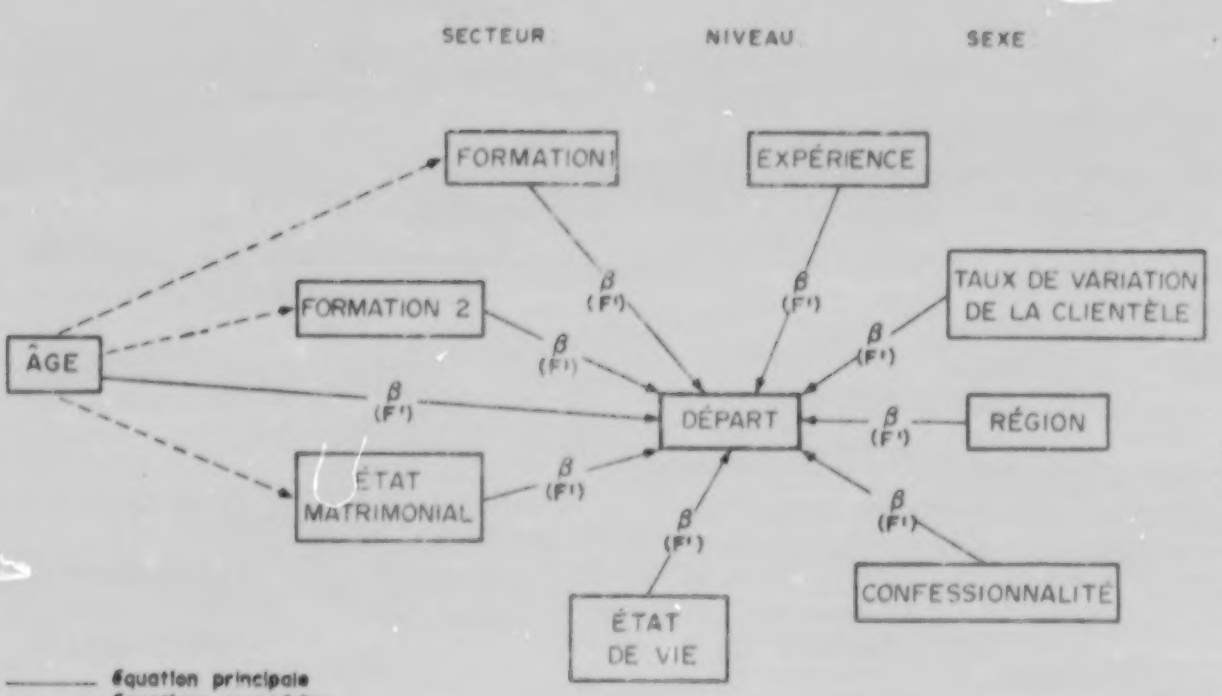
- Certaines variables explicatives déterminent en partie les autres: c'est le cas de l'âge, par exemple, qui agit sur l'état matrimonial et la formation.

Ces remarques nous ont conduit à construire un système graphique de représentation qui tienne compte à la fois de l'équation principale (probabilité de départ en fonction des variables explicatives) et des équations secondaires (liaison des variables explicatives entre elles). Il a la forme suivante:

---

(1) Nous aurions d'ailleurs pu ne produire qu'une seule équation pour tout le système d'éducation, en incluant le secteur, le niveau et le sexe dans les variables qualitatives.

VARIABLES INFLUENÇANT LE DÉPART



avec: n = nombre d'observations
F = test de Snedecor sur l'ensemble de l'équation principale (y compris la constante) (1)
F' = test de Snedecor sur le coefficient de la variable explicative (appelé parfois F "partiel")(1)
β = Coefficient de régression standardisé ou coefficient BETA

(Nota: la constante de l'équation principale est toujours positive)

- Explication sur le coefficient BETA:

A partir d'une équation de régression Y = ax+bz+c,

les coefficients: $a' = a\frac{\sigma(x)}{\sigma(y)}$ $b' = b\frac{\sigma(x)}{\sigma(y)}$

rendent comptent de la relation entre les variables normées

$\frac{x}{\sigma(x)}$ $\frac{y}{\sigma(y)}$ $\frac{z}{\sigma(z)}$ dont les mesures sont exprimées en unité

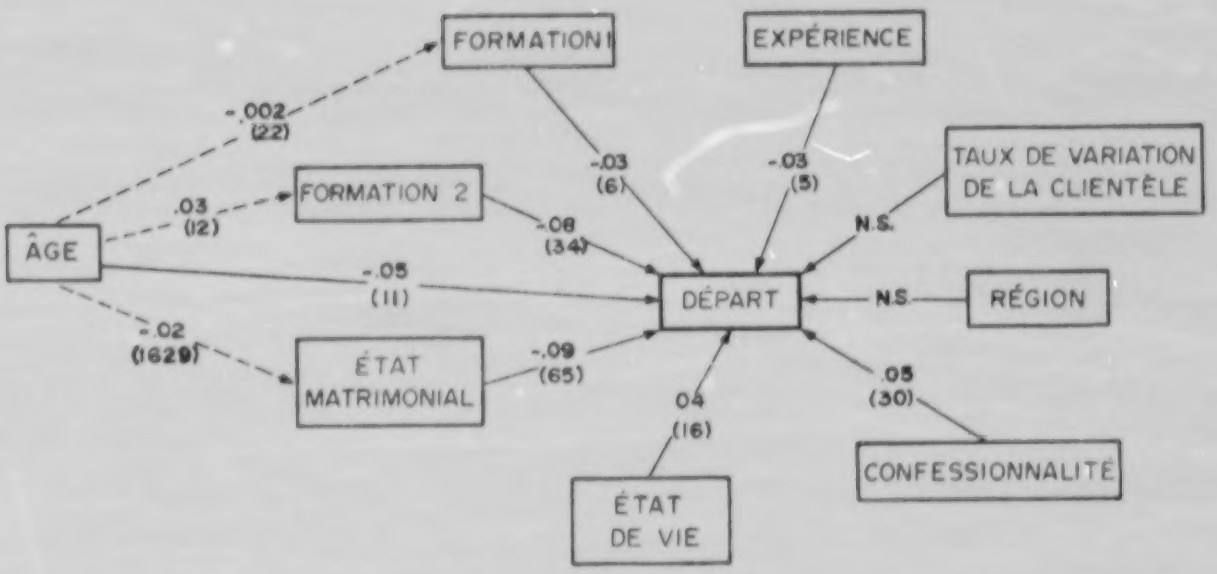
(1) Etant donné le nombre d'observations et le nombre de variables explicatives (qui détermine le nombre de degrés de liberté), il suffit que le F (ou le F') soit égal à 2.8 environ pour être significatif à 99% et à 2 pour 90%.

d'écart-type (procédé de standardisation). La comparaison des coefficients a', b' (coefficients Beta) permet d'estimer l'importance de la liaison entre chaque variable explicative et la variable dépendante: on peut ainsi hiérarchiser les influences, même si les variables explicatives ne sont pas exprimées dans les mêmes unités.

L'application de ces principes donne, surtout pour le secteur public, des résultats acceptables: nous avons donc choisi ce secteur comme illustration, par l'intermédiaire des graphiques no 4, 5, 6, 7, 8, 9, 10 (1).

GRAPHIQUE 4 : VARIABLES INFLUENÇANT LE DÉPART

SECTEUR Public NIVEAU Mat. SEXE Féminin

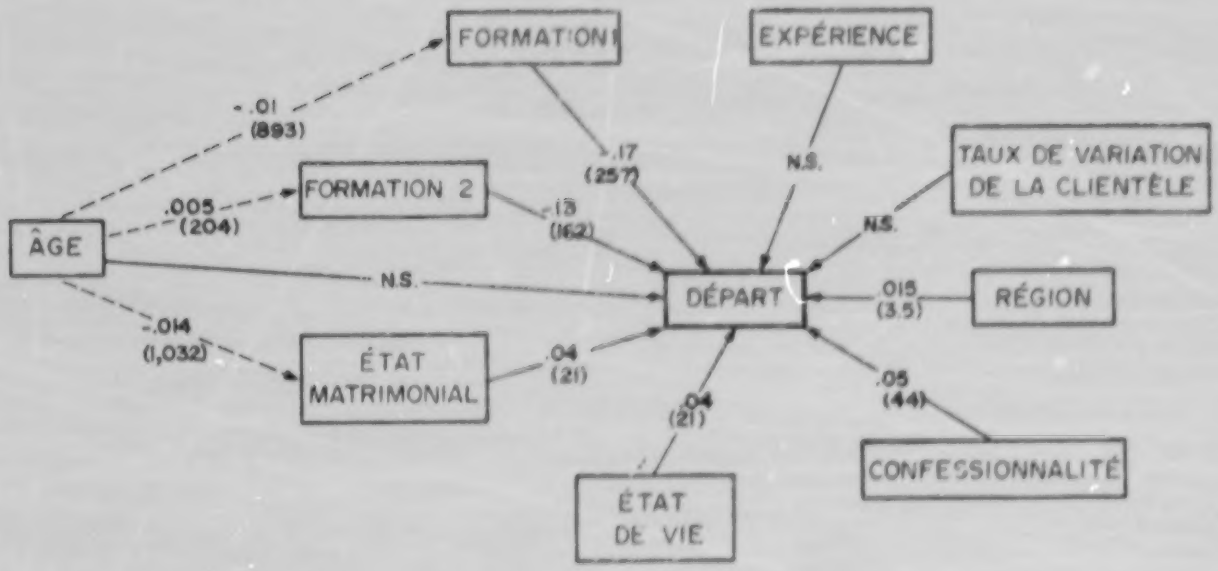


N = 9717 F = 22 N.S. = Non significatif

(1) L'équation principale est obtenue à l'aide de la procédure "stepwise regression" qui facilite la sélection de la "bonne" équation; les équations secondaires proviennent d'une régression linéaire simple (y = ax+b), sans contrôle des autres variables, pour ne pas alourdir la présentation.

GRAPHIQUE 5 : VARIABLES INFLUENÇANT LE DÉPART

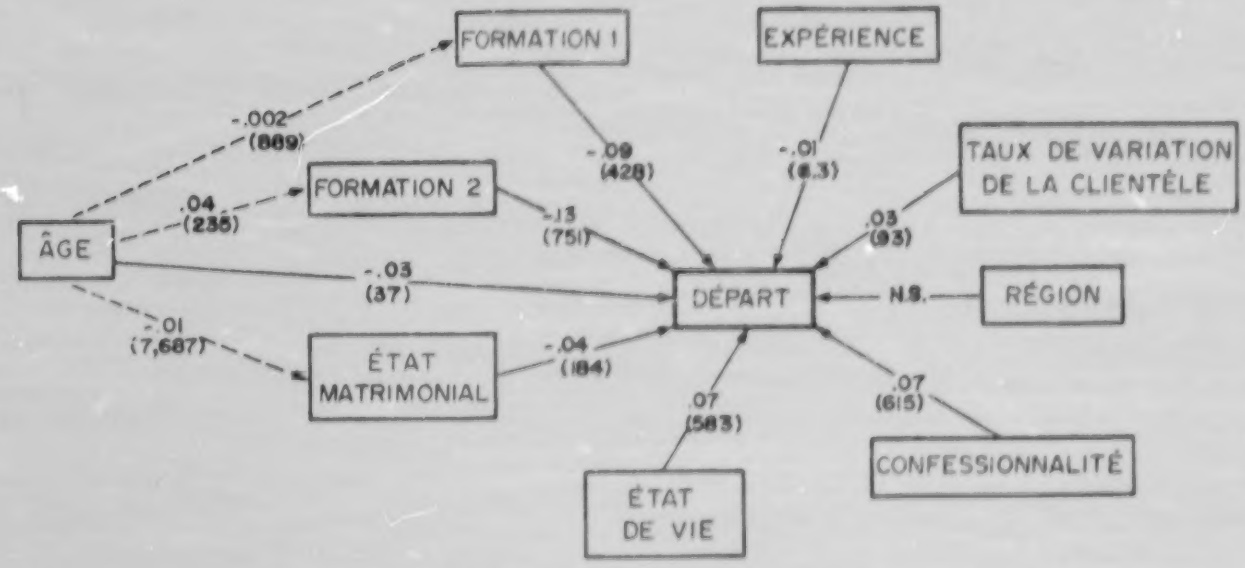
SECTEUR : Public NIVEAU : Élém. SEXE : Masculin



N = 14,765 F = 60.5 N.S = Non significatif

GRAPHIQUE 6 : VARIABLES INFLUENÇANT LE DÉPART

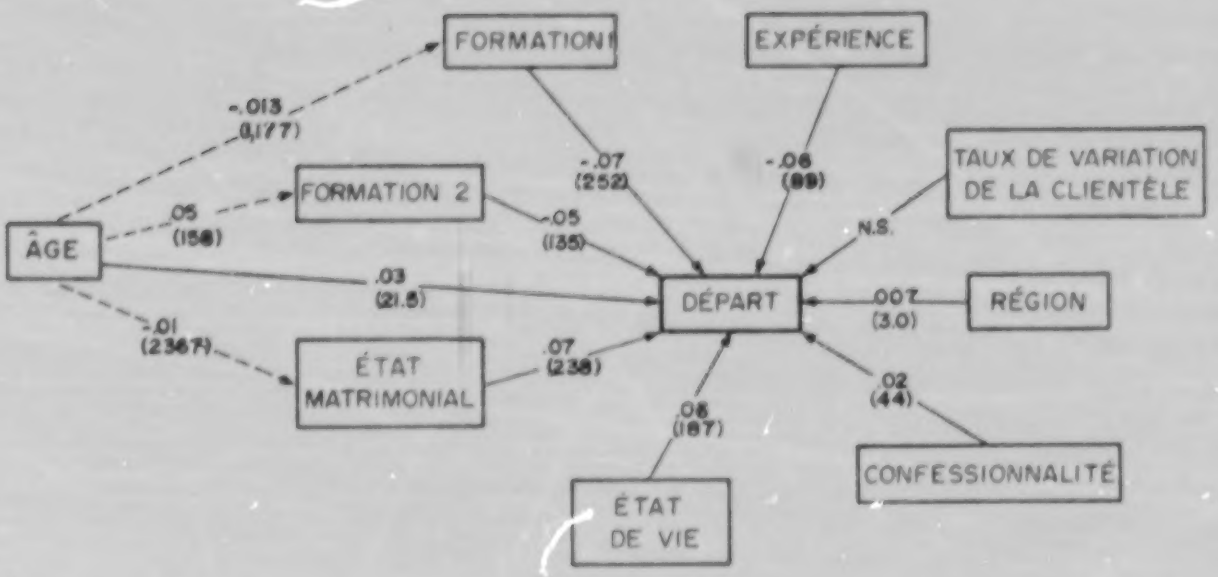
SECTEUR : Public NIVEAU : Élém. SEXE : Féminin



N = 126,251 F = 289.5 N.S. = Non significatif

GRAPHIQUE 7 : VARIABLES INFLUENÇANT LE DÉPART

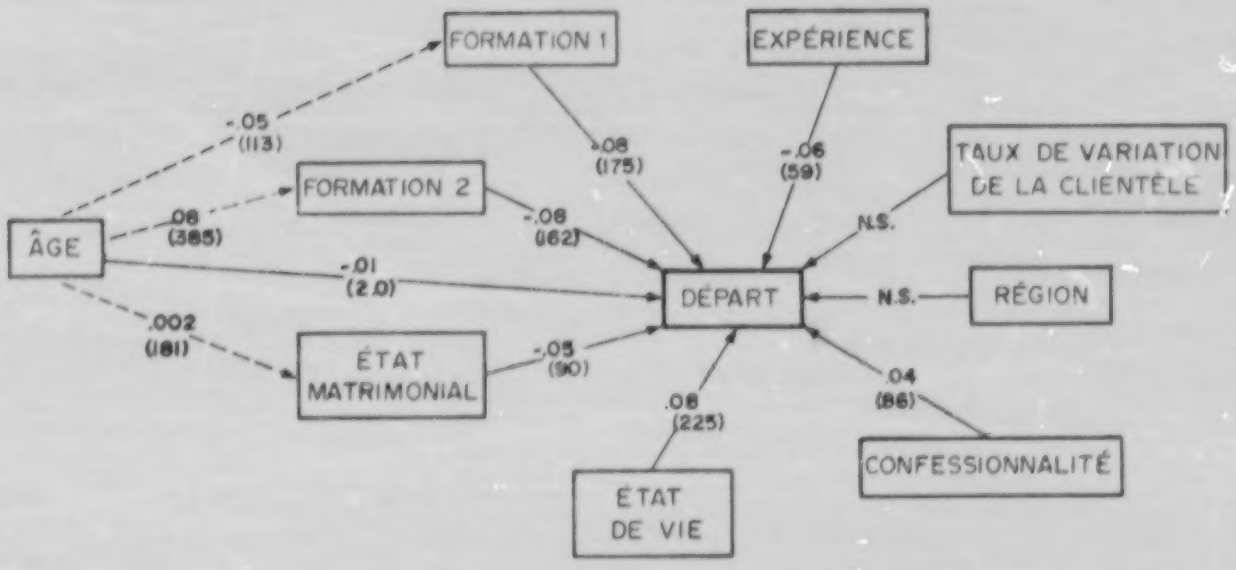
SECTEUR Public NIVEAU Sec. SEXE Masculin



N = 64,225 F = 131.4 N.S = Non significatif

GRAPHIQUE 8 : VARIABLES INFLUENÇANT LE DÉPART

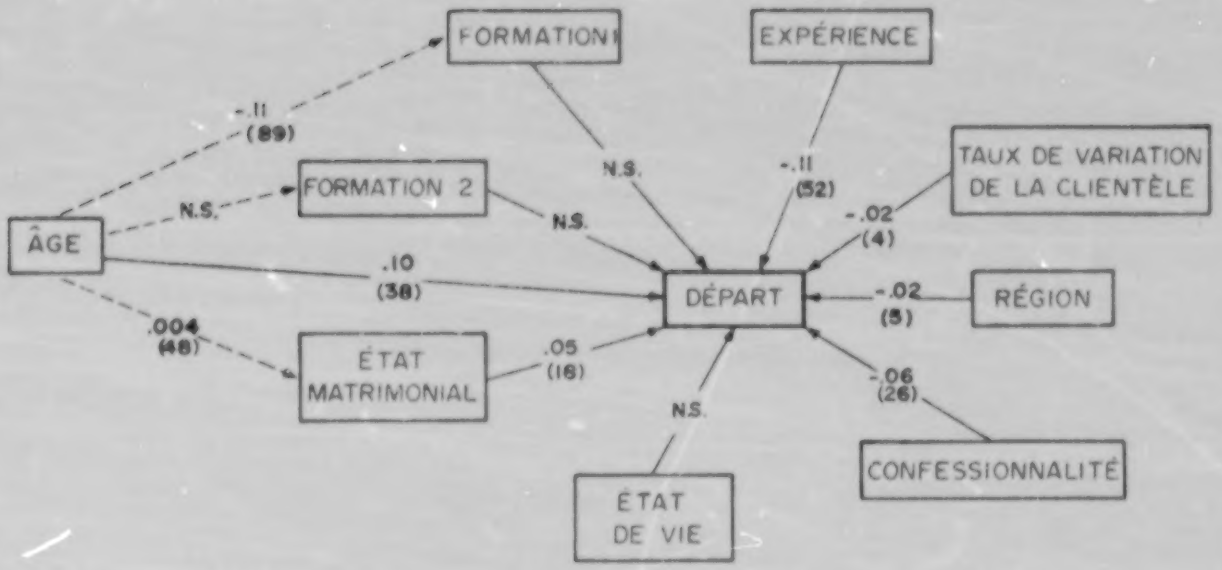
SECTEUR: Public NIVEAU: Sec. SEXE: Féminin



N = 52,317 F = 85 N.S = Non significatif

GRAPHIQUE 9 : VARIABLES INFLUENÇANT LE DÉPART

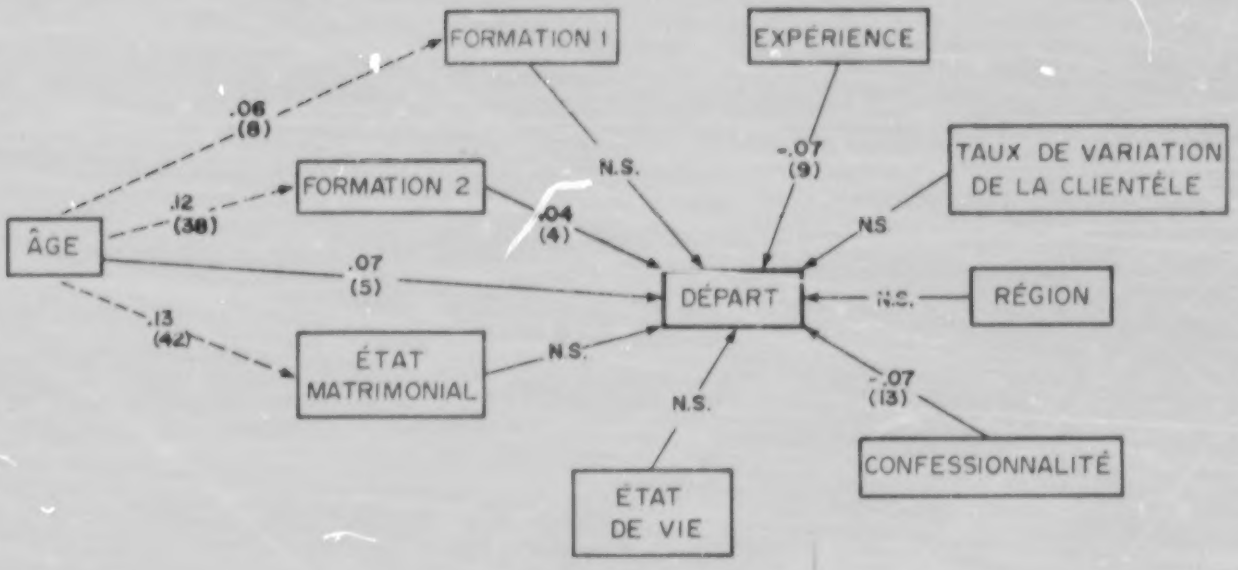
SECTEUR : Public NIVEAU : Collégial SEXE : Masculin



N = 7,284 F = 17.3 N.S = Non significatif

GRAPHIQUE 10 : VARIABLES INFLUENÇANT LE DÉPART

SECTEUR : Public NIVEAU : Collégial SEXE : Féminin



N = 2,393 F = 6.3 N.S = Non significatif

3.4 Interprétation

Le modèle utilisé présente, comme nous l'avons noté, des défauts qui l'empêchent d'être utilisé comme source de prévision. Il permet d'établir néanmoins, toutes choses égales d'ailleurs, l'action de certaines variables.

3.4.1. Les équations secondaires

Ces équations illustrent, sans contrôle des autres variables, l'action de l'âge sur certaines variables explicatives.

- L'âge est négativement lié avec l'état matrimonial, telle que cette variable est définie (1 pour les célibataires) : c'est-à-dire que le nombre des célibataires diminue quand l'âge augmente. Ce qui est logique. La seule exception venant du "Collégial-Féminin" où les religieuses sont à la fois relativement nombreuses, âgées et...célibataires!

- L'âge est négativement lié à la "Formation 1" et positivement lié à la "Formation 2". Résultat intéressant qui démontre l'existence d'une coupure assez nette dans le personnel: les enseignants les plus qualifiés sont concentrés aux âges jeunes. Ce qui est également logique.

3.4.2. L'équation principale

• La "Formation 1" et la "Formation 2" sont, toutes deux, négativement liées avec le "départ": ce qui sous-entend que ce sont les autres catégories de formation (en particulier les "sans diplôme ni brevet") qui partent le plus. Rien de surprenant à cela. Par contre la hiérarchie des effets entre les 2 catégories de formation est intéressante. Elle laisse apparaître, pour les niveaux inférieurs, une distinction selon le sexe: les femmes de formation "moyenne" (Formation 2) sont plus réfractaires au départ que celles qui ont une formation "supérieure" (Formation 1) et inversement pour les hommes.

• L'action de l'âge n'est pas uniforme: le coefficient est négatif pour les femmes jusqu'au secondaire (les jeunes partent donc relativement plus), et devient positif pour le collégial (2 sexes) et le secondaire masculin (les enseignants âgés partent plus). L'explication peut résider dans le fait que ce sont justement ces niveaux, le collégial et le secondaire, qui ont connu le plus de transformations pédagogiques. On verra toutefois, au chapitre 2, qu'une autre technique de régression permet de tirer de l'âge une information plus simple et plus opérationnelle.

• L'état matrimonial laisse apparaître une distinction surprenante entre les sexes: chez les hommes, les célibataires partent plus, alors que c'est l'inverse chez les femmes.

• Le coefficient de l'état de vie est toujours positif, lorsqu'il est significatif: cela signifie, et c'est peut-être surprenant, que les religieux partent plus que les laïcs, les autres variables étant constantes (y compris l'âge).

• La confessionnalité de l'école laisse apparaître une distinction selon le niveau: le coefficient est positif aux niveaux "inférieurs", c'est-à-dire que le départ est plus fréquent dans les écoles protestantes (confessionnalité =1). L'inverse est observé pour le niveau collégial.

• La région, telle que définie, est rarement significative; il en est de même pour le taux de variation de clientèle.

• Enfin l'expérience est toujours négativement liée avec le départ; ce qui confirme le point d'analyse que nous avions soulevé précédemment, à partir du graphique 3.

RAPPEL

- Une forte proportion de l'effectif quitte, chaque année le système de l'éducation: le roulement de personnel est donc important parmi les enseignants.

- Au cours des périodes observées, ce mouvement a atteint son maximum à la fin de 1968-69 ou de 1969-70, et son minimum à la fin de 1970-71.

- La fréquence des départs est plus élevée au secteur privé.

- Le secondaire et le collégial sont plus touchés que l'élémentaire.

- Les hommes partent généralement plus que les femmes.

- Une expérience faible et une formation peu poussée favorisent le départ (au secteur public).

- Les religieux partent plus, toutes choses égales d'ailleurs, que les laïcs (au secteur public).

- La confessionnalité de l'école et l'état matrimonial de l'enseignant ont une influence qui diffère, pour le secteur public, selon le niveau ou le sexe.

- L'influence des grandes agglomérations (Montréal et Québec) est un peu significative, au secteur public.

Chapitre 2 **LE MODELE - DEP'ART -**

Ce modèle, composé d'équations de régression (une pour chaque combinaison du secteur, du niveau et du sexe), utilise l'âge comme seule variable explicative; nous verrons bientôt selon quelles modalités.

Outre le modèle, basé sur les variables dichotomiques, que nous avons décrit dans le chapitre précédent, une autre tentative avait été élaborée; à savoir, bâtir les équations à partir de variables explicatives, grossières sans doute, mais faciles à observer, comme l'âge moyen, l'expérience moyenne, la proportion de célibataires, la variation de clientèle, etc(1)... Au niveau des régressions cette démarche semblait donner des résultats acceptables mais lorsqu'il s'est agi de simuler une prévision, le système a rapidement donné des résultats aberrants. Deux causes à cela: 1° Le nombre d'observations était insuffisant. 2° Certaines variables explicatives, de par leur évolution normale, sortaient de l'espace de régression et ne garantissaient donc plus l'usage de l'équation dans laquelle elles apparaissaient.

Toutefois on pouvait se douter, après avoir pris connaissance des informations contenues dans le chapitre 1, que l'âge serait un indicateur acceptable soit par son influence

---

(1) Cette méthode n'était valable que dans la mesure où elle concernait un modèle à court terme; de plus, l'informatisation du système, en facilitant le remplacement des équations après un supplément d'information, permettait d'atténuer l'effet du nombre restreint d'observations.

directe sur la probabilité de départ, soit par sa liaison avec une variable-relais telle que l'expérience, l'état matrimonial ou la formation; avec, en plus, l'avantage d'être une variable simple, traditionnelle dans l'analyse démographique et dont la distribution dans l'effectif est assez régulière. Encore fallait-il vérifier cette hypothèse de travail.

1. Estimation des probabilités de départ selon l'âge

1.1 La technique de régression

Nous aurons déjà noté, en examinant le graphique 2, que la comparaison des structures par âge, départs versus l'effectif de base, suggérait que la fréquence relative (1) du départ était plus faible au centre de l'intervalle d'âge qu'aux extrêmes.

L'observation empirique des distributions confirme, comme l'indique le graphique 11, cette hypothèse. La fonction mathématique susceptible d'exprimer cette relation devra donc avoir la forme d'une courbe en U, plus ou moins déformée selon le cas.

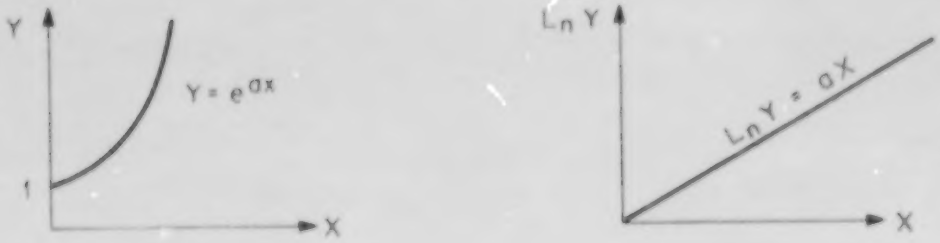
L'utilisation de la régression dite "curviligne", nécessaire à priori pour approximer de telles fonctions, est en

---

(1) la fréquence relative est, en général, calculée sur un effectif standardisé de 100 (%); En la divisant par 100 (donc en la ramenant à un effectif standardisé de 1), on peut assimiler cette fréquence relative à une probabilité (variant entre 0 et 1).

général très délicate, tant du point de vue de la théorie (1) que de la pratique (car les programmes disponibles sur ordinateur suivent, la plupart du temps, la technique de régression "linéaire") Mais la pratique de l'anamorphose (2) permet souvent de tourner cet obstacle. On appelle "anamorphose" toute transformation de variable qui permet, sans modifier la relation, d'exprimer une fonction non-linéaire sous une forme linéaire (3). L'utilisation du Logarithme est l'exemple traditionnel d'une telle pratique; en effet les deux expressions suivantes sont équivalentes:

$$Y = e^{\alpha x} \iff \text{Ln } Y = \alpha x \text{Ln } e = \alpha x$$



On voit que la première est une fonction exponentielle, alors que la seconde est linéaire. Inversement on peut avoir l'équivalence suivante:

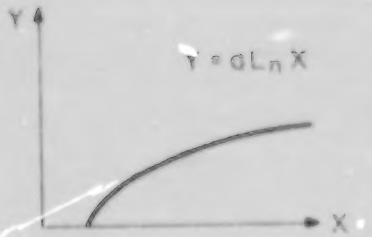
---

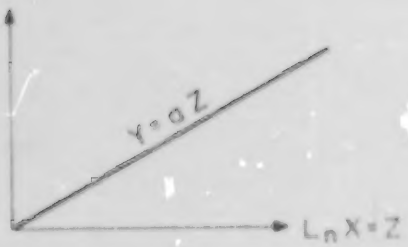
(1) cf. Draper & Smith: "Applied Regression Analysis", chap. 10 ("an introduction to nonlinear estimation"), Wiley & Sons, 1966.

(2) cf. J. Dubos: "Liaisons stochastiques en économie", chap. 3., Dunod, 1971.

(3) Selon le définition la plus générale, l'anamorphose est "la transformation géométrique d'une figure suivant certaines règles".

$Y = a \operatorname{Ln} X \longleftrightarrow Y = aZ$ avec $Z = \operatorname{Ln} X$



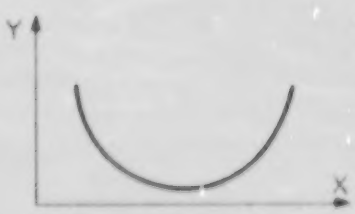


Le changement de variable permet donc d'exprimer sous une forme linéaire, une fonction logarithmique. La généralisation de ce système, au niveau de la variable dépendante (Y) ou de la variable explicative (X) ou des deux à la fois, va donc permettre d'exprimer toute une famille de fonctions sous une forme linéaire et autoriser ainsi l'emploi des méthodes traditionnelles de régression linéaire. Il en est de même des fonctions polynomiales, très utiles pour notre étude;

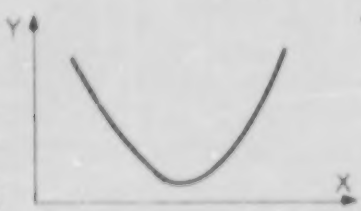
En posant $X^2 = U$, $X^3 = V$, ecc... on a:

$$Y = a + bX + cX^2 + dX^3 \longleftrightarrow Y = a + bX + cU + dV$$

L'application de ces principes conduit, entre autres, à la régression des fonctions suivantes:



$Y = a + bX + cX^2$



$\operatorname{Ln} Y = a + bX + cX^2$

$(Y = e^{(a+bX+cX^2)})$

$Y = a + b \operatorname{Ln} X + c \operatorname{Ln} X^2$

Pour sélectionner la meilleure régression, nous avons utilisé le test dit de Snedecor (ou test de F) et la comparaison des coefficients de détermination (R^2): étant donné le nombre d'observations utilisé dans les régressions, le F de l'équation doit, au minimum, être égal à 5 pour assurer une significativité à 99%[1], avec deux variables explicatives[2]. On peut alors s'assurer, comme le montre le tableau suivant, que le recours aux "anamorphoses" améliore considérablement la qualité des estimations:

Comparaison entre diverses régressions

Secteur public, Secondaire, Masculin

Nombre d'observations: 205

Variable dépendante dans la régression	Variables explicatives dans la régression	Type de fonction estimée	F	R^2
Prob. de départ	Age	$Y=aX+b$	19.4	.09
Log.Prob.de départ	Age	$Y=e^{ax}e^{b}$	24.2	.10
Prob. de départ	Age, $(Age)^2$	$Y=aX+bX^2+c$	119.8	.54
Log.Prob.de départ	Age, $(Age)^2$	$Y=e^{(ax+bx^2+c)}$	144.7	.59
Prob. de départ	Log. Age	$Y=aL_nX+b$	6.7	.03
Log.Prob.de départ	Log. Age	$Y=X^{a}e^{b}$	8.7	.04
Prob. de départ	Log.Age,$(Log.Age)^2$	$Y=aL_nX+b(L_nX)^2+c$	112.7	.53
Log.Prob.de départ	Log.Age,$(Log.Age)^2$	$Y=e(aL_nX+b(L_nX)^2+c)$	149.3	.60

(1) On trouve des tables de signification du F (à 99%, 95% ou 75%) dans les manuels de statistiques; par exemple: DUBOS, op. cit., p. 471.

(2) Le R^2, quant à lui, doit être aussi proche que possible de 1.

1.2 Les résultats

Les équations ainsi obtenues ne dépendent que de l'âge et ne comportent aucun indice de calendrier, ce qui est acceptable pour un modèle à court terme. Il est néanmoins possible d'introduire un certain effet temporel en modifiant les périodes d'observation à partir desquelles est réalisée la régression. C'est ainsi que, disposant de 4 périodes d'observations (1), nous avons décidé d'utiliser 3 simulations:

- Simulation no 1: la régression est calculée sur l'ensemble des observations disponibles.

- Simulation no2: la régression est calculée sur trois périodes, après exclusion de la période 1968-69 / 1969-70, période qui semble présenter des probabilités de départ plus élevées; on fait donc implicitement l'hypothèse que l'année exclue est un "accident", aux causes diverses, sans influence durable sur le phénomène des départs.

- Simulation no 3: la régression est affectuée sur la dernière période seulement (1970-71 / 1971-72); on essaie ainsi, même si la réduction du nombre des ob-

(1) de la fin de 1967/68 à celle de 1970/71.

servations peut réduire la "qualité" de certaines équations, de retracer l'évolution du phénomène en tenant compte, essentiellement, de la dernière information disponible (1).

Les équations de probabilité de départ par âge, qui découlent de ces 3 simulations, sont données dans les tableaux 14,15 et 16, tandis que le graphique 11 illustre, dans le cas du "Public-Secondaire-Masculin", les écarts entre les observations et l'équation de la simulation no 1. Il faut noter que nous avons exclu les "Maternelle-Masculin" car l'effectif y est négligeable.

On voit qu'en général les estimations sont moins bonnes au secteur privé qu'au public, et, à l'intérieur des niveaux, moins bonnes à la maternelle et au collégial qu'à l'élémentaire et qu'au secondaire. Outre les variations possibles de comportement, cet écart s'explique par le fait que les fréquences empiriques de départ-qui sont à la base de la régression-sont d'autant plus instables que l'effectif est faible.

Il faut, toutefois, se rappeler deux choses:

1° Pour la qualité de la prévision, l'erreur doit être faible là où l'effectif est important; ce qui semble être le cas.

---

(1) procédé qui permet d'intégrer, s'il existe, l'effet temporel le plus récent.

PROB DE DÉPART

GRAPHIQUE 11 : COMPARAISON DES PROBABILITÉS DE DÉPART

SECTEUR PUBLIC , SECONDAIRE , MASCULIN

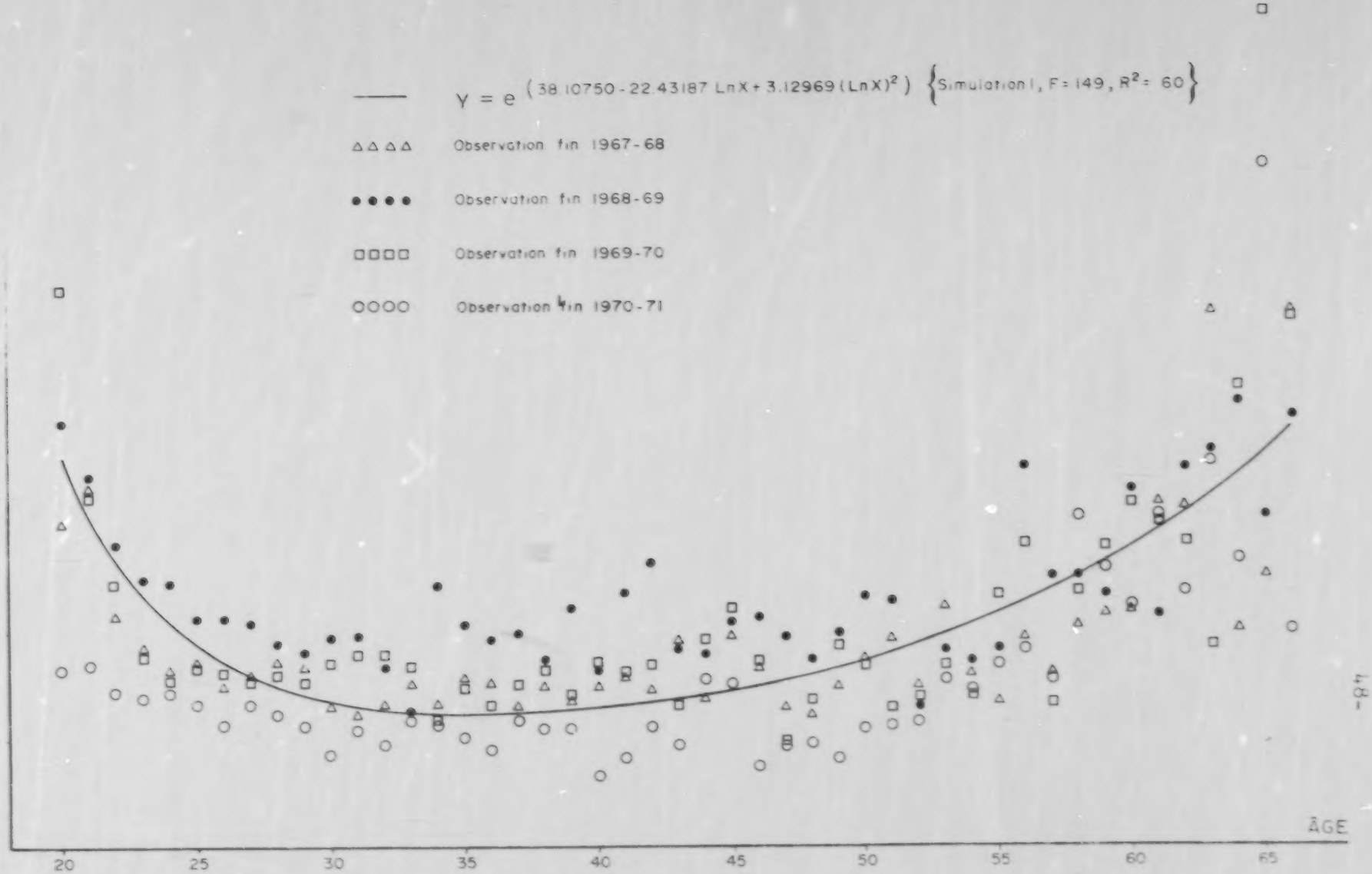


TABLEAU 14: Equations des probabilités de départ - Simulation no 1

avec { y= Probabilité de départ à un âge donné
x= Age
n= Nombre d'observations

Catégorie de référence			Equation de Y (*)	F	R^2	n
SECTEUR PUBLIC	SEXE MASCULIN	Elémentaire	$.97856 - .04595X + .00059X^2$	87.1	.49	185
		Secondaire	$e^{(38.10730 - 22.43187 L_nX + 3.12869 (L_nX)^3)}$	149.3	.60	205
		Collégial	$.78528 - .0372X + .00048X^2$	34.5	.28	184
	SEXE FEMININ	Maternelle	$.81295 - .04060X + .00037X^2$	101.3	.57	155
		Elémentaire	$e^{(1.17558 - .18597X + .00236X^2)}$	284.3	.73	212
		Secondaire	$1.07230 - .04765X + .00060X^2$	131.6	.56	209
		Collégial	$10.16240 - 5.60830 L_nX + .79828 (L_nX)^2$	20.7	.20	165
SECTEUR PRIVE	SEXE MASCULIN	Elémentaire	$e^{(12.26971 - 7.87411 L_nX + 1.16598 (L_nX)^2)}$	12.1	.18	114
		Secondaire	$10.37994 - 5.62269 L_nX + .77839 (L_nX)^2$	45.2	.32	198
		Collégial	$.80468 - .02676X + .00036X^2$	32.9	.28	174
	SEXE FEMININ	Maternelle	$1.24354 - .04566X + .00062X^2$	13.6	.18	125
		Elémentaire	$9.49461 - 5.10405 L_nX + .70522 (L_nX)^2$	42.0	.30	199
		Secondaire	$.96217 - .03213X + .00037X^2$	30.3	.23	203
		Collégial	$1.27773 - .04119X + .00047X^2$	20.8	.20	173

(*) Tous les coefficients de régression sont significatifs à 99%

TABLEAU 15: Equations des probabilités de départ - Simulation no 2 (exclusion de 1968-69 / 1969-70)

avec { y= Probabilité de départ à un âge donné
x= Âge
n= Nombre d'observations

Catégorie de référence			Equation de Y (*)	F	R^2	n
SECTEUR PUBLIC	SEXE MASCULIN	Elémentaire	$11.63604 - 6.49744 \ln X + .91195 (\ln X)^2$	59.3	.47	138
		Secondaire	$e^{(1.05412 - .15706 X + .00194 X^2)}$	116.9	.61	153
		Collégial	$12.03426 - 6.63866 \ln X + .94079 (\ln X)^2$	35.1	.34	138
	SEXE FEMININ	Maternelle	$.86952 - .04424 X + .00061 X^2$	90.3	.61	118
		Elémentaire	$e^{(1.42267 - .20433 X + .00260 X^2)}$	288.7	.79	159
		Secondaire	$e^{(1.19950 - .18948 X + .00196 X^2)}$	137.3	.64	157
		Collégial	$10.16434 - 5.63027 \ln X + .80426 (\ln X)^2$	15.2	.20	123
SECTEUR PRIVE	SEXE MASCULIN	Elémentaire	$8.39648 - 4.70931 \ln X + .69107 (\ln X)^2$	11.4	.23	80
		Secondaire	$10.09844 - 5.44754 \ln X + .75043 (\ln X)^2$	35.1	.33	147
		Collégial	$.85914 - .08206 X + .00044 X^2$	30.8	.34	124
	SEXE FEMININ	Maternelle	$10.62793 - 5.81577 \ln X + .82546 (\ln X)^2$	8.7	.17	87
		Elémentaire	$10.49493 - 5.66093 \ln X + .77906 (\ln X)^2$	38.8	.35	147
		Secondaire	$1.01032 - .03551 X + .00041 X^2$	27.8	.27	153
		Collégial	$1.32161 - .04287 X + .00049 X^2$	13.8	.19	124

(*) Tous les coefficients de régression sont significatif à 99%

TABLEAU 16: Equations des probabilités de départ - Simulation no 3 (Période 1969-70/1970-71 seule)

avec { y= Probabilité de départ à un âge donné
x= Age
n= Nombre d'observations

Catégorie de référence			Equation de Y	F	R^2	n
SECTEUR PUBLIC	SEXE MASCULIN	Elémentaire	$.67462 - .03325 X + .00046 X^2$	21.3	.51	43
		Secondaire	$e^{(-.03575 - .12000 X + .00159 X^2)}$	38.5	.62	50
		Collégial	$.91807 - .049096 X + .00053 X^2$	41.5	.67	44
	SEXE FEMININ	Maternelle	$1.04814 - .05991 X + .00072 X^2$	62.1	.77	40
		Elémentaire	$e^{(1.59318 - .22492 X + .00288 X^2)}$	143.8	.85	52
		Secondaire	$e^{(2.04939 - .20923 X + .00280 X^2)}$	69.4	.74	52
		Collégial	$.77262 - .03042 X + .00042 X^2$	10.6	.37	39
SECTEUR PRIVE	SEXE MASCULIN	Elémentaire	$12.77609 - 7.23805 L_n X + 1.05306 (L_n X)^2$	6.2	.38	23
		Secondaire	$.84680 - .03331 X + .00042 X^2$	12.3	.37	45
		Collégial	$1.36197 - .06142 X + .00080 X^2$	13.9	.47	34
	SEXE FEMININ	Maternelle	$1.31666 - .05439 X + .00076 X^2$	4.1 (1)	.26	26
		Elémentaire	$12.43228 - 6.69147 L_n X + .91132 (L_n X)^2$	22.3	.51	46
		Secondaire	$.81616 - .03136 X + .00039 X^2$	12.7	.35	51
		Collégial	$2.09202 - .08549 X + .00102 X^2$	12.3	.43	35

(1) Significatif à 95%; les autres équations sont significatives à 99%

2° L'objet de la prévision est la nombre total de départs et non les départs à un âge donné: un écart à ce niveau peut être compensé et ne pas altérer le résultat final.

Enfin les tableaux 17 et 18, comparent, à certains âges, les probabilités obtenues des diverses simulations. Elles confirment, en général, l'analyse qui a été faite au chapitre 1.

- Ces équations ne sont pas définitives; comme nous le verrons au chapitre suivant, elles doivent être recalculées après tout nouvel apport d'information. On pourra tester, également, des fonctions polynomiales du 3ème degré [$y=ax+bx^2+cx^3+d$, $L_n Y= ax+bx^2+cx^3+d$, $Y= aL_nx+b(L_nx)^2+c(L_nx)^3+ d$] qui pourraient donner de bons résultats dans certaines catégories de référence.

2. Approximation de la structure par âge

La connaissance des probabilités de départ selon l'âge nous oblige, puisque nous faisons une prévision du nombre absolu de départs, à estimer la distribution par âge, des effectifs; et cela pour une année donnée. Disposant de cette distribution (ou structure), nous en déduirons l'effectif à un âge donné, puis, faisant intervenir la probabilité de

TABLEAU 17: Probabilité de départ à certains âges selon les différentes simulations

(Secteur public)

Sexe	Niveau	Simulation	25 ans	35 ans	45 ans	55 ans	65 ans
MASCULIN	ELE.	N° 1	.200	.094	.106	.237	.486
		N° 2	.170	.063	.117	.243	.404
		N° 3	.131	.074	.109	.237	.457
	SEC.	N° 1	.188	.124	.145	.217	.369
		N° 2	.190	.127	.124	.180	.383
		N° 3	.130	.101	.109	.161	.327
	COL.	N° 1	.355	.271	.283	.391	.595
		N° 2	.349	.252	.320	.459	.632
		N° 3	.225	.134	.148	.286	.495
FEMININ	MAT.	N° 1	.154	.090	.140	.304	.582
		N° 2	.145	.068	.114	.282	.571
		N° 3	.150	.043	.080	.261	.586
	ELE.	N° 1	.135	.087	.089	.147	.390
		N° 2	.127	.079	.081	.142	.417
		N° 3	.107	.064	.067	.127	.437
	SEC.	N° 1	.256	.139	.143	.266	.510
		N° 2	.215	.143	.140	.204	.440
		N° 3	.198	.109	.010	.150	.372
	COL.	N° 1	.387	.321	.390	.517	.672
		N° 2	.374	.313	.386	.517	.676
		N° 3	.275	.222	.254	.370	.570

TABLEAU 18: Probabilité de départ à certains âges selon les différentes simulations (secteur privé)

Sexe	Niveau	Simulation	25 ans	35 ans	45 ans	55 ans	65 ans
MASCULIN	ELE.	No 1	.370	.373	.445	.571	.754
		No 2	.398	.389	.484	.622	.780
		No 3	.388	.353	.483	.682	.912
	SEC.	No 1	.346	.228	.256	.348	.472
		No 2	.339	.216	.236	.319	.435
		No 3	.276	.195	.198	.285	.456
	COL.	No 1	.361	.309	.329	.422	.586
		No 2	.332	.276	.307	.427	.634
		No 3	.326	.192	.218	.404	.749
FEMININ	MAT.	No 1	.489	.405	.444	.608	.895
		No 2	.460	.385	.451	.578	.747
		No 3	.431	.344	.408	.624	.992
	ELE.	No 1	.351	.237	.255	.334	.442
		No 2	.345	.216	.235	.320	.439
		No 3	.335	.161	.165	.251	.379
	SEC.	No 1	.390	.291	.265	.314	.437
		No 2	.379	.270	.243	.298	.434
		No 3	.276	.196	.195	.271	.425
	COL.	No 1	.542	.412	.376	.434	.586
		No 2	.556	.421	.385	.446	.605
		No 3	.592	.349	.310	.475	.845

départ à cet âge, le nombre des départs qu'il fournira: le nombre total des départs sera bien entendu la somme de ces départs, en faisant varier l'âge.

Comme nous l'avait montré le graphique 2; la distribution de l'effectif selon l'âge est une courbe "en cloche", unimodale (ou quasi-unimodale si l'effectif est faible), très dissymétrique vers la gauche (avec, en plus, un mode inférieur à la moyenne), plus ou moins leptokurtique ("cloche" de forme aiguë) selon le cas. Un certain nombre de fonctions mathématiques peuvent représenter de telles distributions:

- la fonction dite "Gamma incomplète" (1): $f(x) = \frac{1}{\Gamma(\vartheta)} e^{-x} x^{\vartheta-1}$ (avec $\vartheta > 1$) ; mais cette fonction présente une contrainte rigide selon laquelle la différence entre la moyenne et le mode doit toujours être égale à 1: ce qui est inapplicable à notre cas.

- la dérivée de la fonction de Gompertz (2): $f(x) = KA^{B^{(x-x_0)}} L_n A B^{(x-x_0)} L_n B$. Cette fonction comporte malheureusement, elle aussi, une contrainte: elle inclut obligatoirement, quels que soient les paramètres, 37% de la distribution entre le début de l'intervalle

---

(1) cf. Calot: "Cours de calcul des probabilités", chap. 13, Dunod, 1967.

(2) cf. Murphy et Nagnur: "A Gompertz fit that fits: applications to Canadian fertility patterns", Demography, vol.9, No 1, Février 1972.

d'âges et le mode. Ce qui est beaucoup trop élevé pour nos distributions.

- La fonction "type 1" de Pearson et la fonction "type transition no3" de Pearson, que nous avons utilisées.

Toutes ces fonctions prennent, dans notre cas, la forme de distributions de densité de probabilité. En effet, la structure par âge "empirique" est généralement calculée en pourcentage, c'est-à-dire par rapport à un effectif standardisé de 100; il suffit de ramener cet effectif à 1, pour s'assurer que les conditions d'une loi de densité de probabilité (fonctions variant entre 0 et 1, fonction de répartition égale à 1 sur l'ensemble de l'intervalle) soient réunies.

2.1 Fonction de Pearson, type "transition no3" (ou "Pearson type 3")

Pour une information complète sur le système de courbes de Pearson, et notamment la Pearson type 3, on pourra consulter: W.P. Elderton, "Frequency curves and correlation", Harren Press, 1er ed: 1903, dernière ed: 1953.

L'ensemble du système de Pearson est généré par l'équation différentielle suivante:

$$\frac{df(x)}{dx} = \frac{f(x)(x+a)}{b_0+b_1x+bx^2+\ldots}$$

où le dénominateur est un développement de MacLaurin, d'un degré choisi.

Plus particulièrement la fonction "type 3" présente, avec l'origine placée au mode, l'équation et les caractéristiques suivantes:

$$f(x) = f_0 \left(1 + \frac{x - MOD}{a}\right)^{\gamma a} e^{-\gamma (x - MOD)}$$

$M_3 = 2\,(MOY - MOD)\,VAR$

$\gamma = (2\,VAR) / M_3$

$a = \dfrac{VAR}{MOY - MOD} - (MOY - MOD)$

$p = a\,\gamma$

$f_0 = \dfrac{1}{a} \dfrac{p^{(p+1)}}{e^p\,\Gamma(p+1)}$

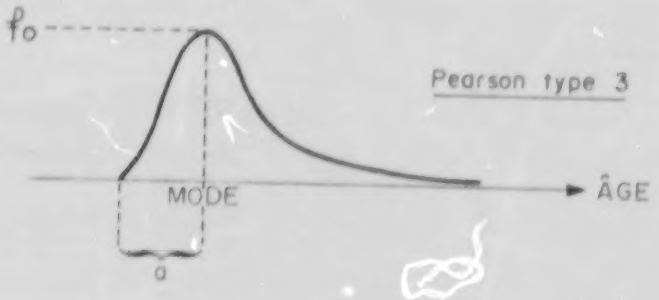
avec

- γ, P paramètres
- M_3 moment centré d'ordre 3
- MOY âge moyen
- MOD âge modal
- VAR variance
- a différence entre l'âge modal et la borne inférieure de l'intervalle d'âge.
- f_0 densité de probabilité à l'âge modal (maximum)
- Γ fonction Gamma complète.

En utilisant la méthode dite " des moments", c'est-à-dire en remplaçant dans les équations l'âge moyen, l'âge modal et la variance par leurs valeurs observées ou prévues, on disposera donc d'une approximation mathématique de la structure par âge de l'effectif.

Il faut noter dans le cas de la Pearson 3 que, le paramètre (a) étant déterminé à l'intérieur du système d'équa-

tions, la fonction détermine elle-même la borne inférieure (âge modal - a) de l'intervalle d'âge; la borne supérieure, elle, est infinie, si bien que la fonction a la forme suivante:



2.2 Fonction de Pearson "type 1"

En plus de W.P. Elderton, op.cit., on pourra consulter les références suivantes:

S. Mitra: "The pattern of age-specific fertility rates", Demography, 1967, no2, pp. 894-906.

Shryock, Siegel & autres: "The methods and materials of demography" vol.2, chap.22, p.693, U.S. Bureau of Census, 1971.

S. Mitra & A.Romaniuk: "Pearsonian type I curve and its fertility projection protentials", Demography, vol. 10, no3, Août 1973.

Comme on peut le constater en lisant la bibliographie, ce type de fonction a surtout été employé dans les prévisions

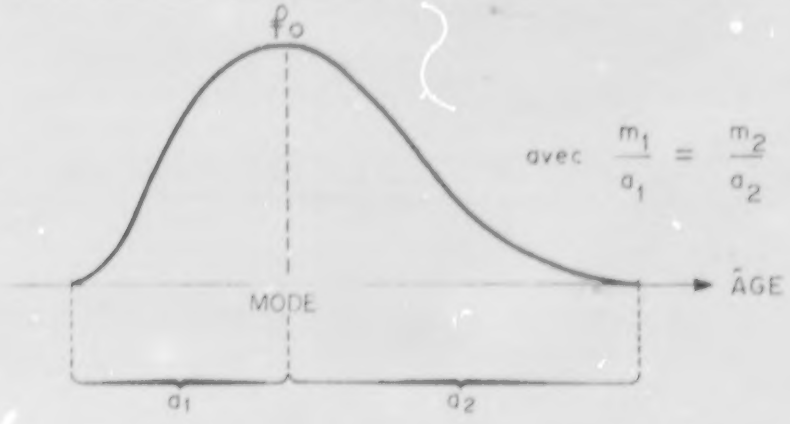
de taux de fécondité, mais peut être utilisé dans n'importe quel problème du genre de celui qui nous préoccupe.

Dans sa forme la plus complète, les paramètres de l'équation sont calculés à partir des 4 premiers moments centrés(1), mais au niveau de la prévision, le 3ème et le 4ème moment (qui servent à déterminer les indices d'asymétrie et de kurtose)(2) sont très difficiles à interpréter; ce qui n'est pas le cas pour le 1er moment (la moyenne) ou le 2ème (la variance).

Nous avons donc suivi la méthode développée par S. Mitra en 1967, méthode qui consiste à estimer les paramètres à partir des seules moyenne et variance.

La fonction est de la forme:

$$f(x) = f_0 \left(1 + \frac{X - MOD}{a_1}\right)^{m_1} \left(1 - \frac{X - MOD}{a_2}\right)^{m_2}$$



(1) cf. Elderton, op. cit.

(2) kurtose est synonyme d'applatissement (cf. l'appendice statistique)

L'avantage de cette fonction réside dans le fait que les paramètres a_1 et a_2 peuvent être fixés, après examen des séries empiriques. Reste la détermination de m_1 et m_2 : Mitra utilise la procédure suivante(1):

$$m_1+m_2 = \frac{(MOY-MOD+a_1)(a_2-MOY+MOD)}{VAR} - 3$$

$$m_1 = \frac{m_1+m_2+2}{a_1+a_2}(MOY-MOD+a_1)-1 \longrightarrow m_2 = (m_1+m_2)-m_1$$

Nous avons rencontré une difficulté lors de l'utilisation de cette procédure: étant donné les valeurs des moments et du mode dans nos distributions, la condition $\frac{m_1}{a_1} = \frac{m_2}{a_2}$ n'était pas toujours respectée. Ce qui avait pour conséquence, en particulier, de ne pas assurer que la densité de probabilité maximum se trouvât à l'âge modal désigné. Nous avons alors décidé de "forcer" cette condition et d'obtenir les paramètres m_1 et m_2 de la manière suivante:

$$m_1+m_2 = \frac{(MOY-MOD+a_1)(a_2-MOY+MOD)}{VAR} - 3$$

$$m_2 = \frac{(m_1+m_2)}{1+\frac{a_1}{a_2}} \qquad m_1 = m_2 \cdot \frac{a_1}{a_2}$$

Ce qui n'est donc qu'une approximation de la forme "pure" de la Pearson type 1, mais permet de résoudre notre problème.

---

(1) extrapolée de Elderton, op. cit., p. 64; cf. S. Mitra op. cit. (1967)

4.3 Les résultats

Les deux procédures (Pearson 3 et Pearson 1) ont été incluses dans le programme de prévision; l'utilisateur peut donc choisir l'une ou l'autre. Quelques points pratiques doivent être, cependant, précisés:

- intégration dans l'intervalle d'âge: la fonction de densité de probabilité nous donne une valeur à un âge exact (32.47 ans, par exemple) alors que nos distributions empiriques sont calculées en âge révolu (ceux qui ont "25 ans" ont, en fait, entre 25 et 26 ans). Si nous voulons avoir une distribution en âge révolu, ce qui est nécessaire pour faciliter l'intervention des équations de probabilité de départ, il nous faut intégrer la fonction de Pearson entre deux âges:

$$F(x) = \int_x^{x+1} f(x)\,dx$$

La fonction de Pearson, quelle que soit son numéro, n'ayant pas une équation particulièrement simple, la tâche peut sembler ingrate. Il est d'usage, heureusement, d'avoir recours dans un cas comme celui-ci à l'approximation de Simpson:

$$F(x) = \int_x^{x+1} f(x)\,dx = \frac{1}{6}\left(f(x) + 4 f(x+0.5) + f(x+1) \right)$$

Cette approximation donne d'excellents résultats dans un intervalle aussi court.

- Contrôle d'effectif: dans la Pearson 3, et même dans la Pearson 1, la procédure informatique ne couvre pas l'ensemble de l'intervalle de définition et ne répartit donc pas 100% de l'effectif. Pour éviter cette perte de personnel, il est nécessaire d'avoir recours à une "deuxième boucle" et la procédure est alors la suivante (en language P.L.1.):

```
/*1ère BOUCLE*/
DØ X=B TØ BB,
   f(x)=
   F(x)=
   TØT.F=TØT.F+F(x),
END;

/*2ème BOUCLE*/
DØ X=B TØ BB,
   F(x)=F(x)/TØT.F,
END,
```

- Il peut arriver qu'avec les indications qui lui sont données (âge moyen, âge modal, variance), la fonction de Pearson 3 ne puisse fonctionner convenablement, soit parce qu'elle génère un (a) négatif, soit qu'elle prenne la forme d'une distribution en J.

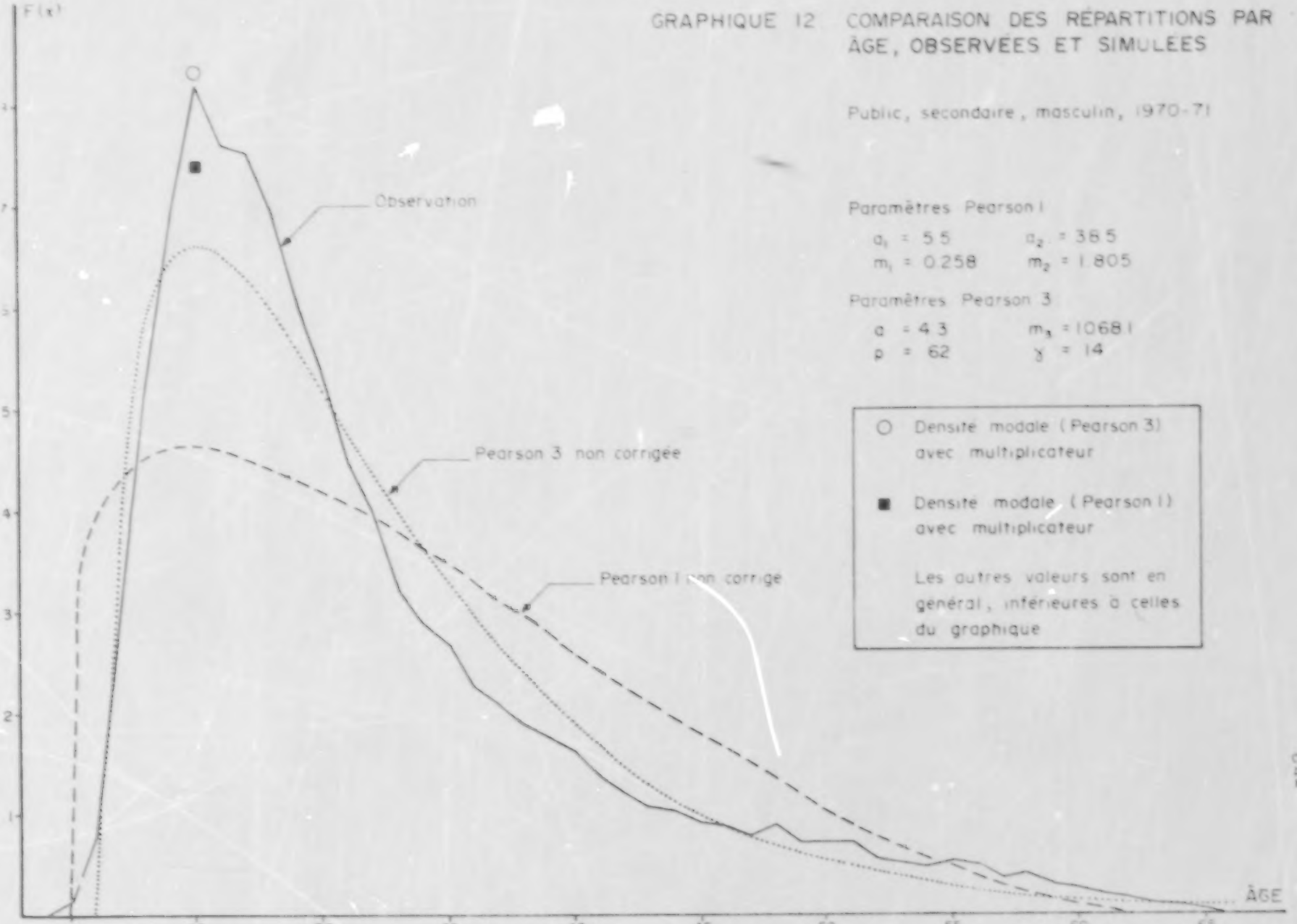
Le programme exécute alors, automatiquement, un changement de procédure en faveur de la Pearson 1.

- Multiplicateur modal: lors de l'expérimentation, il est apparu, surtout pour la Pearson 1, que la densité de probabilité maximum (densité à l'âge modal) était nettement insuffisante, comparée avec l'observation. Un "multiplicateur", propre à chacune des fonctions, a alors été calculé pour rapprocher ces 2 valeurs: il a été appliqué en entier à la densité modale des 2 fonctions et, partiellement, aux densités immédiatement proches de la densité modale, pour la Pearson 1. Cette pratique a pour conséquence de faire répartir plus de 100% du personnel; excès qui est corrigé par l'utilisation de la " 2e boucle".

On voit donc que, pour résoudre nos problèmes particuliers, nous avons dû faire subir quelques déformations à nos fonctions théoriques. Donner une évaluation comparée de leur utilisation est assez délicat puisque, notre prévision portant sur 14 catégories (secteur, niveau, sexe) et 11 années (de 1967-68 à 1977-78), notre système génère, par exemple, 154 distributions de Pearson 1 (un peu moins pour la Pearson 3, à cause des transferts automatiques), toutes avec des paramètres différents... Néanmoins, le graphique 12 illustre

GRAPHIQUE 12 COMPARAISON DES RÉPARTITIONS PAR ÂGE, OBSERVÉES ET SIMULÉES

Public, secondaire, masculin, 1970-71



Paramètres Pearson I

$a_1 = 5.5$ $a_2 = 38.5$

$m_1 = 0.258$ $m_2 = 1.805$

Paramètres Pearson 3

$\alpha = 4.3$ $m_3 = 10681$

$\beta = 62$ $\gamma = 14$

○ Densité modale (Pearson 3) avec multiplicateur

■ Densité modale (Pearson I) avec multiplicateur

Les autres valeurs sont en général, inférieures à celles du graphique

la démarche, en prenant comme référence le "Public, Secondaire, Masculin" de 1970-71, tandis que les tableaux 19 et 20, en prenant comme point de comparaison le taux brut de départ obtenu par la combinaison d'une fonction de distribution et d'une équation de probabilité, permettant de "tester" le résultat global sur les années passées (1). Evidemment, le fait d'utiliser la même équation de probabilité de départ pour toutes les années (alors même que la répartition par âge ne peut se modifier qu'assez lentement) amortit considérablement les variations du taux brut par rapport aux observations. D'autre part, le fait d'utiliser toutes les années comme base de la régression (simulation no 1) sur-estime le départ lorsqu'on envisage une année "faible" (1970-71, par exemple) et le sous-estime lorsqu'on est en présence d'une année "forte" (1968-69 ou 1969-70); et cela d'autant plus que les variations sont importantes.

Les deux fonctions donnent, en fait, des résultats à peu près équivalents ; néanmoins, en examinant ces résultats(2), nous en avons réparti l'utilisation de la manière suivante:

---

(1) En fait, le multiplicateur, s'il permet d'avoir des distributions plus proches de la réalité, ne modifie qu'assez peu le taux brut de départ: celui-ci dépend surtout des équations utilisées.

(2) Nous n'avons pas eu le temps de mettre au point un programme qui, en calculant "l'écart-type" de l'ajustement, aurait permis un jugement plus précis.

TABLEAU 19: Comparaison des taux bruts de départ obtenus avec les fonctions de Pearson

(Probabilité de départ: Simulation no 1)

De fin 1967-68 à fin 1970-71. Secteur public

			MATERNELLE	ELEMENTAIRE	SECONDAIRE	COLLEGIAL
SEXE MASCULIN	1967-68	OB.	--	.129	.167	.330
		P1.	--	.162	.165	.314
		P3.	--	.177	.173	.320
	1968-69	OB.	--	.158	.214	.290
		P1.	--	.158	.165	.315
		P3.	--	.170	.173	.315
	1969-70	OB.	--	.143	.174	.437
		P1.	--	.161	.165	.313
		P3.	--	.177	.176	.313
	1970-71	OB.	--	.116	.124	.172
		P1.	--	.161	.165	.312
		P3.	--	.177	.163	.313
SEXE FEMININ	1967-68	OB.	.139	.143	.212	.387
		P1.	.165	.131	.210	.410
		P3.	.165	.132	.219	.372
	1968-69	OB.	.182	.178	.266	.383
		P1.	.153	.126	.210	.400
		P3.	.160	.131	.215	.373
	1969-70	OB.	.146	.139	.213	.429
		P1.	.148	.123	.208	.388
		P3.	.155	.127	.205	.368
	1970-71	OB.	.128	.111	.155	.263
		P1.	.144	.120	.207	.384
		P3.	.155	.123	.210	.374

(OB.: observation, P1: Pearson 1, P3: Pearson 3)

TABLEAU 22a: Comparaison des taux bruts de départ obtenus avec les fonctions de Pearson

(Probabilité de départ: Simulation no 1)

De fin 1967-68 à fin 1970-71. Secteur privé

	NIVEAU		MATERNELLE	ELEMENTAIRE	SECONDAIRE	COLLEGIAL
SEXE MASCULIN	1967-68	OB.	--	.264	.239	.298
		P1.	--	.453	.297	.343
		P3.	--	.403	.295	.333
	1968-69	OB.	--	.394	.348	.440
		P1.	--	.440	.303	.342
		P3.	--	.398	.287	.334
	1969-70	OB.	--	.361	.356	.386
		P1.	--	.426	.300	.343
		P3.	--	.382	.275	.335
	1970-71	OB.	--	.276	.233	.211
		P1.	--	.408	.302	.341
		P3.	--	.393	.277	.332
SEXE FEMININ	1967-68	OB.	.347	.294	.320	.402
		P1.	.488	.314	.335	.456
		P3.	.454	.290	--	--
	1968-69	OB.	.439	.342	.416	.463
		P1.	.486	.315	.334	.453
		P3.	.463	.296	--	.447
	1969-70	OB.	.416	.344	.373	.547
		P1.	.492	.318	.333	.457
		P3.	.467	.303	.317	.474
	1970-71	OB.	.361	.257	.254	.442
		P1.	.497	.315	.334	.454
		P3.	.472	.293	.320	.462

(OB.: observation, P1: Pearson 1, P3: Pearson 3)

Pearson 1 corrigée

Public-Maternelle-Féminin
Public-Elémentaire-Masculin
Public-Collégial-Masculin
Privé-Secondaire-Masculin
Privé-Secondaire-Féminin
Privé-Collégial-Féminin

Pearson 3 corrigée

Public-Elémentaire-Féminin
Public-Secondaire-Masculin
Public-Secondaire-Féminin
Public-Collégial-Féminin
Privé-Maternelle-Féminin
Privé-Elémentaire-Masculin
Privé-Elémentaire-Féminin
Privé-Collégial-Masculin

- Il faut noter, enfin, qu'il est possible de mener l'étude par groupe d'âges: il faut, pour cela, soit intégrer les équations de probabilité sur 5 ans, soit recalculer une régression sur les observations empiriques, classées par groupe d'âges (1). De plus, l'intégration de la fonction de Pearson doit se faire dans un intervalle d'âge de 5 ans (2).

(1) Ce qui réduit considérablement le nombre d'observations.

(2) Il est préférable, dans ce cas, d'utiliser 2 fois la formule de Simpson (cf: Mitra, op.cit., 1967)

Chapitre 3 RESULTAT DES SIMULATIONS

Après cette description méthodologique plutôt aride, nous allons passer au résultat. Dans un premier temps nous décrirons le système informatique, puis nous analyserons l'initialisation de différentes variables nécessaires à la prévision, enfin nous présenterons et comparerons le nombre de départs prévu par les simulations.

1. Informatisation du système de calcul

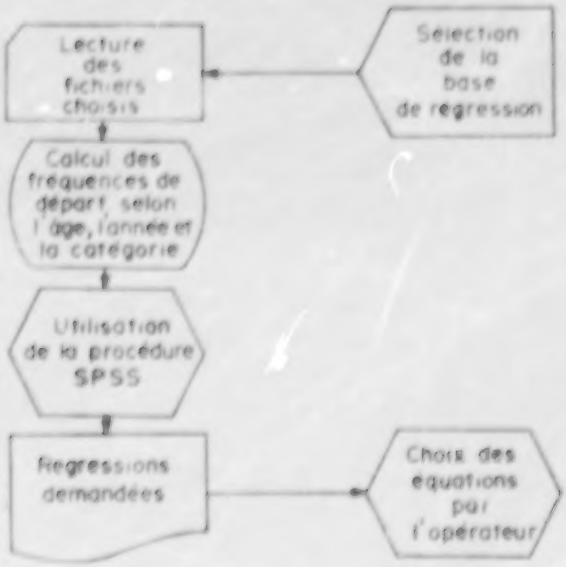
Nous avons déjà précisé dans l'introduction que l'informatisation était, en fait, le principal intérêt pratique de ce travail, et qu'en particulier, ce procédé devait faciliter la révision des prévisions. La logique du programme (qui est divisé en 2 parties) est illustrée à la page suivante. Tout renseignement plus précis, se rapportant à la programmation par exemple, pourra être obtenu de la D.G.P.

2. Initialisation des variables nécessaires à la prévision

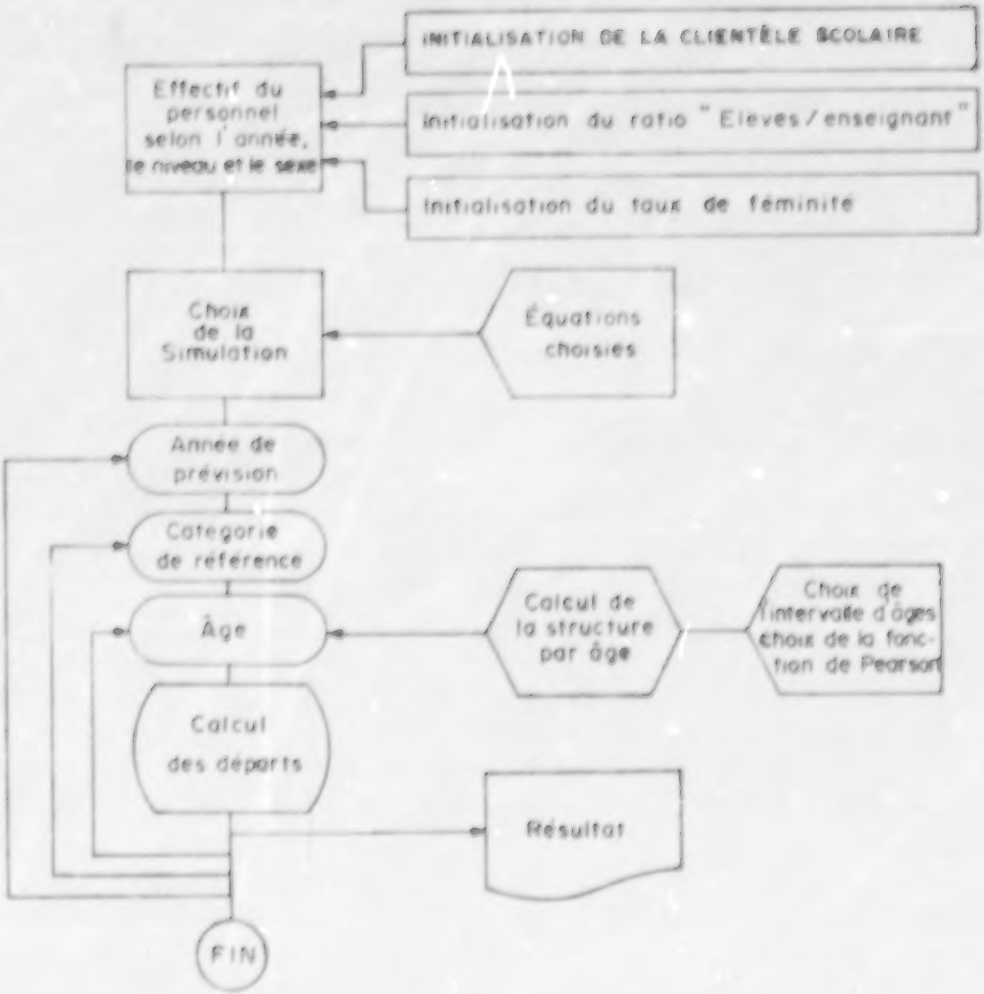
Pour estimer le nombre de départs parmi les enseignants, il est nécessaire de connaître l'effectif de base probable de ces enseignants, une année donnée. Un tel effectif est directement lié, par le jeu du ratio "Elèves / enseignant",

LOGIQUE DU PROGRAMME DE PRÉVISION

1ère Partie :


Lecture des fichiers choisis
Sélection de la base de regression
Calcul des fréquences de départ selon l'âge, l'année et la catégorie
Utilisation de la procédure SPSS
Regressions demandées
Choix des équations par l'opérateur

2ème Partie


INITIALISATION DE LA CLIENTÈLE SCOLAIRE
Effectif du personnel selon l'année, le niveau et le sexe
Initialisation du ratio " Eleves / enseignant "
Initialisation du taux de féminité
Choix de la Simulation
Équations choisies
Année de prévision
Catégorie de référence
Âge
Calcul de la structure par âge
Choix de l'intervalle d'âges choix de la fonction de Pearsat
Calcul des départs
Résultat
FIN

à la clientèle scolaire prévue pour l'année en question; on voit ainsi que notre simulation dépend d'une autre prévision, celle des clientèles. Enfin, il faut répartir les enseignants selon le sexe, ce qui nécessite l'intervention du taux de féminité (Nombre de femmes / Effectif total). Ce sont là ce que nous appellerons des variables "structurelles", par opposition aux caractéristiques d'âge, nécessaires pour l'utilisation des fonctions de distribution, dont nous discuterons par la suite.

2.1. Les variables "structurelles"

- La clientèle scolaire a fait l'objet d'une estimation jusqu'en 1977-78, par le groupe "Démographie scolaire" de la D.G.P.(1): ce qui explique le choix de cette année-là, comme date limite de notre prévision. Nous avons utilisé l'hypothèse no 1 (hypothèse "faible") de Michel Amyot; les effectifs qui en découlent sont données dans le tableau 21.

- Le ratio "Eleves-enseignant"qui détermine l'encadrement de la clientèle, a été extrapolé en tenant compte de sa tendance passée et les valeurs "officielles" inscrites dans les conventions collectives (2).

---

(1) cf. Michel Amyot: "Estimation de la clientèle scolaire, de 1972-73 à 1977-78". M.E.Q., D.G.P., Doc. 9.15, déc. 73.

(2) Ces valeurs officielles, pour le secteur public, sont de 40/1, à la maternelle, de 26/1 à l'élémentaire, de 17/1 au secondaire, et de 15/1 au collégial.

TABLEAU 21: Hypothèse d'évolution des clientèles scolaires(1).

		1971/72	1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
SECTEUR PUBLIC	MATERNELLE	100,881	95,431	92,000	89,900	87,950	85,300	73,750
	ELEMENTAIRE	836,473	756,397	717,550	691,550	663,500	637,200	613,350
	SECONDAIRE	632,005	672,696	654,650	616,050	584,500	543,950	488,750
	COLLEGIAL	78,844	88,877	97,250	105,600	107,250	111,350	122,200
SECTEUR PRIVE	MATERNELLE	3,506	3,583	3,100	3,000	3,000	2,900	2,500
	ELEMENTAIRE	21,437	20,305	18,250	16,500	15,125	14,000	13,200
	SECONDAIRE	40,355	48,031	54,000	59,000	63,500	65,400	65,400
	COLLEGIAL	11,197	11,306	11,435	11,835	12,335	13,540	15,345

(1) - M. Amyot: "Estimation de la clientèle scolaire, de 1972-73 à 1977-78" (hypothèse no 1). D.G.P., Doc. 9.15, décembre 1973.

Les valeurs retenues pour le secteur public sont inférieures à ces valeurs officielles parce qu'elles tiennent compte de l'enfance inadaptée qui nécessite un encadrement plus poussé. Le secteur privé rejoint le secteur public, lorsque son ratio était plus élevé, dans la mesure où il y est obligé pour garder sa clientèle. L'évolution des ratios retenue est illustrée dans les graphique 13 et 14.

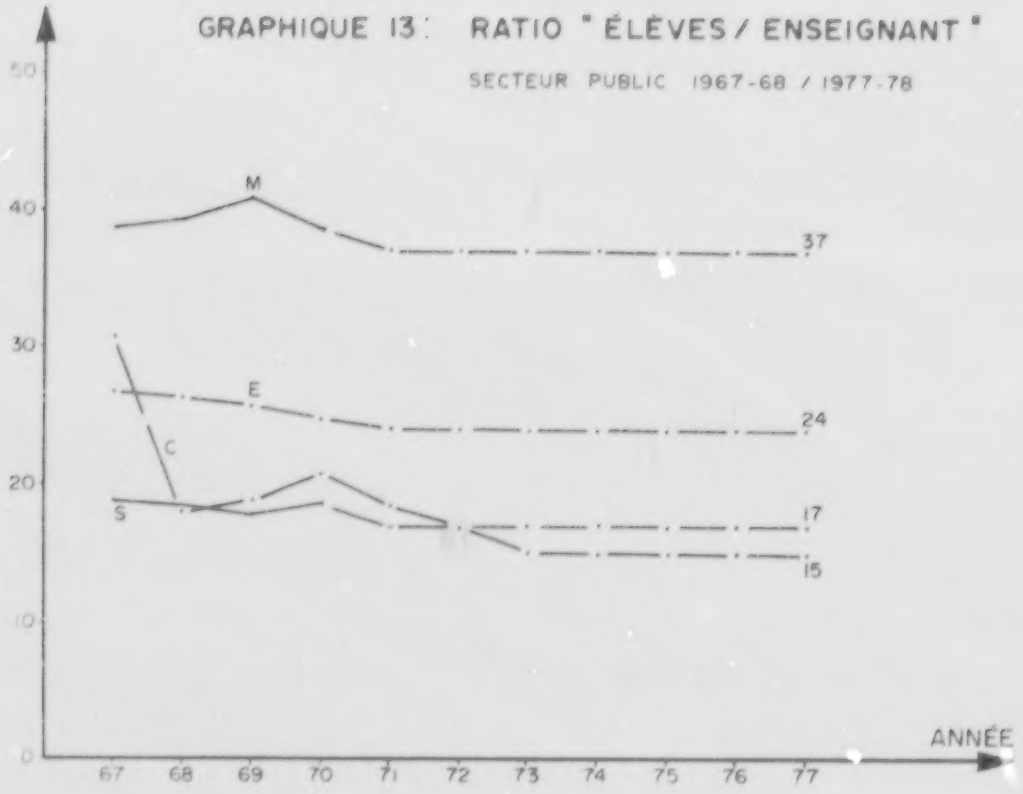
- Le taux de féminité est extrapolé des tendances passées, bien que le faible nombre de données (4 observations) pousse souvent à utiliser une hypothèse de constance. Les valeurs retenues pourront être analysées grâce aux graphiques 15 et 16 (1).

- L'ensemble de ces hypothèses nous permet de simuler l'évolution de l'effectif du personnel selon le secteur, le niveau et le sexe: ces effectifs sont regroupés dans les tableaux 22 et 23.

(1) Les hommes ont été exclus de la maternelle (taux de féminité égal à 1), parce que leurs effectifs y étaient insuffisants.

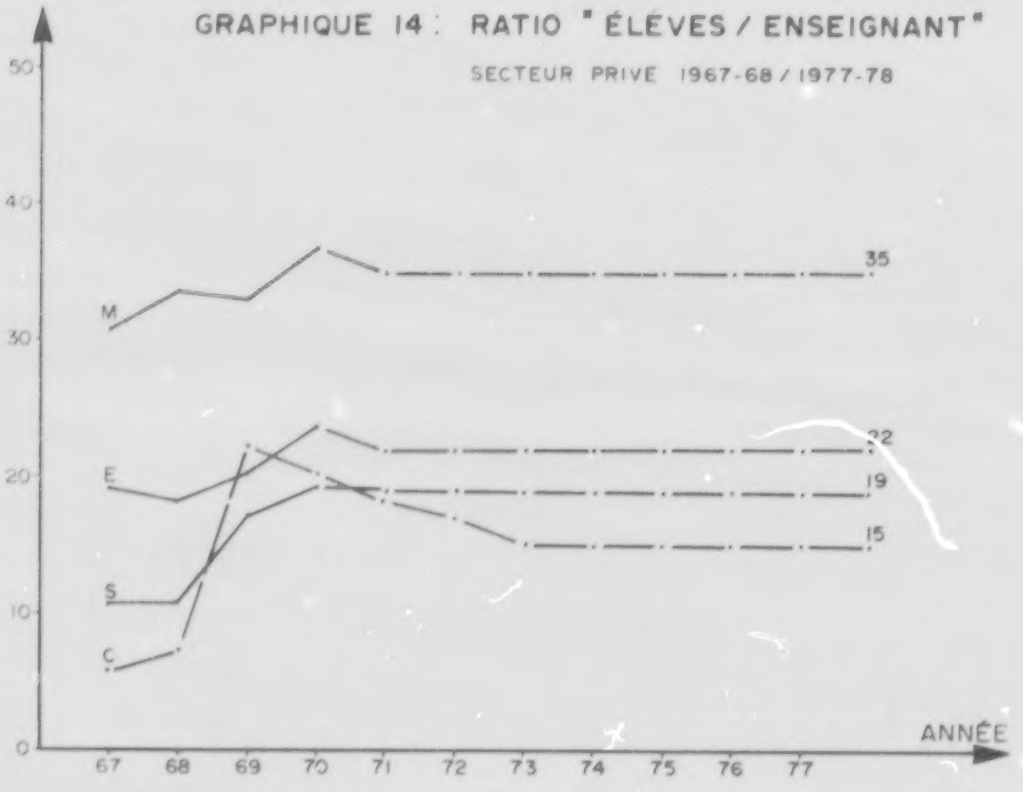
GRAPHIQUE 13 : RATIO " ÉLÈVES / ENSEIGNANT "

SECTEUR PUBLIC 1967-68 / 1977-78



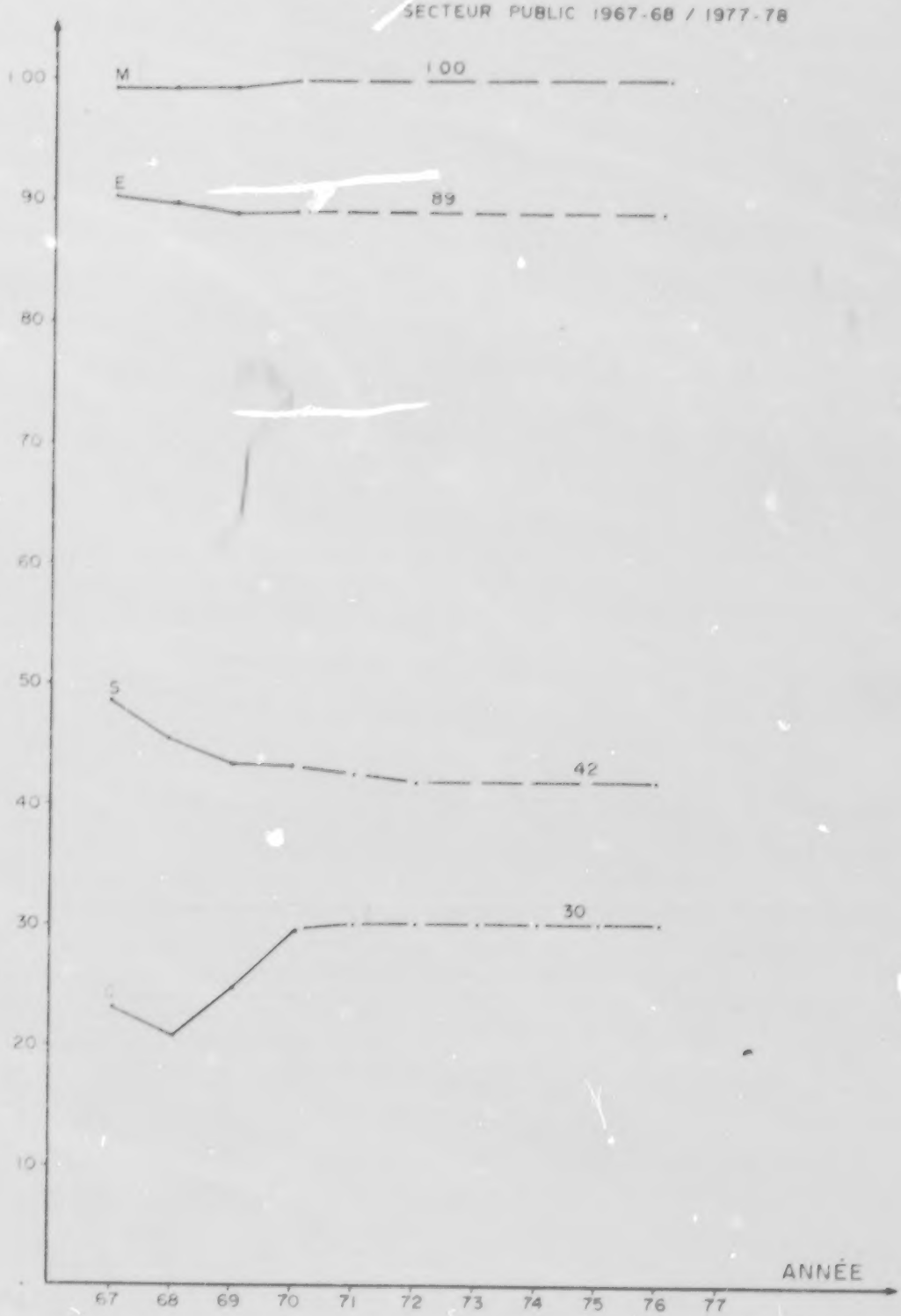
GRAPHIQUE 14 : RATIO " ÉLÈVES / ENSEIGNANT "

SECTEUR PRIVE 1967-68 / 1977-78



GRAPHIQUE 15 : TAUX DE FÉMINITÉ DU PERSONNEL

SECTEUR PUBLIC 1967-68 / 1977-78



GRAPHIQUE 16 TAUX DE FÉMINITÉ DU PERSONNEL

SECTEUR PRIVÉ , 1967-68 / 1977-78

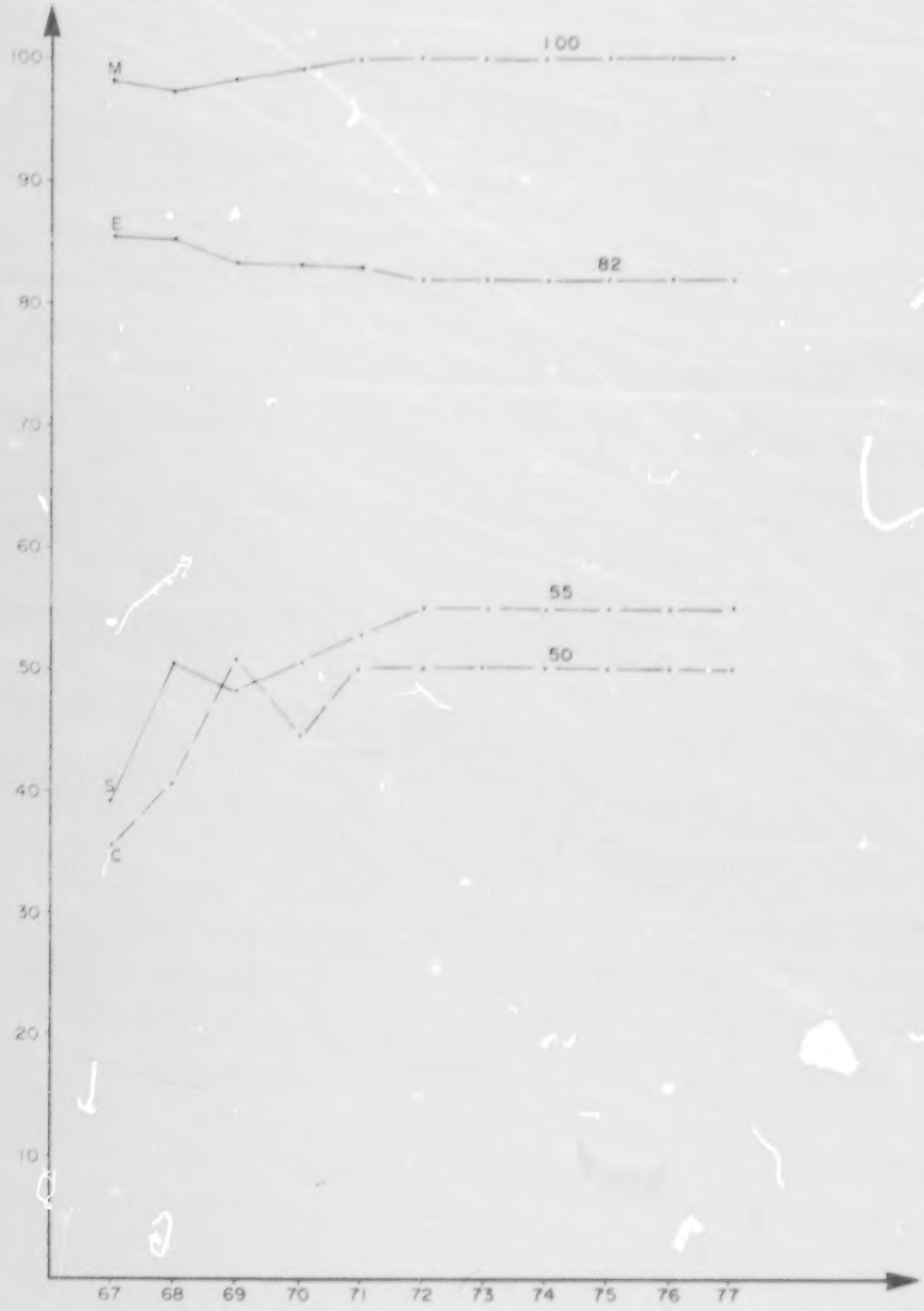


TABLEAU 72: Evolution des effectifs du personnel enseignant

Sexe masculin

		1971/72	1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
SECTEUR PUBLIC	MATERNELLE	----	----	----	----	----	----	----
	ELEMENTAIRE	3,833	3,466	3,288	3,169	3,041	2,920	2,811
	SECONDAIRE	21,376	22,950	22,335	21,018	19,941	18,558	16,675
	COLLEGIAL	2,983	3,959	4,538	4,928	5,005	5,196	5,702
SECTEUR PRIVE	MATERNELLE	----	----	----	----	----	----	----
	ELEMENTAIRE	165	166	149	135	123	114	108
	SECONDAIRE	998	1,137	1,278	1,397	1,503	1,548	1,548
	COLLEGIAL	302	332	381	394	411	451	511

TABLEAU 23: Evolution des effectifs du personnel enseignant

Sexe féminin

		1971/72	1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
SECTEUR PUBLIC	MATERNELLE	2,726	2,579	2,486	2,429	2,377	2,305	1,993
	ELEMENTAIRE	31,019	28,049	26,609	25,644	24,604	23,629	22,745
	SECONDAIRE	15,800	16,619	16,173	15,220	14,440	13,438	12,075
	COLLEGIAL	1,278	1,568	1,944	2,112	2,145	2,226	2,443
SECTEUR PRIVE	MATERNELLE	100	102	88	88	85	82	71
	ELEMENTAIRE	808	756	680	616	563	521	492
	SECONDAIRE	1,125	1,396	1,563	1,777	1,838	1,893	1,893
	COLLEGIAL	302	332	381	394	411	451	511

4.4 Les variables touchant l'âge

Nous avons vu que les fonctions de Pearson utilisées pour répartir l'effectif entre les âges nécessitaient au moins 3 paramètres: l'âge moyen, la variance autour de l'âge moyen et de l'âge modal. Nous y avons ajouté, lorsque c'était nécessaire, l'indication de l'intervalle d'âge (âge de début, âge de fin). En fait les variations de ces différents paramètres, s'ils "déforment" la distribution selon l'âge, ont un effet marginal sur l'évolution du taux brut de départ. On l'avait déjà soupçonné aux tableaux 19 et 20, cela se confirme avec les tableaux 24 et 25; aussi nous n'allons pas infliger au lecteur des chiffres supplémentaires, nous contentant d'énoncer les différentes hypothèses qui ont servi de base à la simulation des variables concernées (1).

- l'âge moyen du secteur public augmente à la maternelle, à l'élémentaire et au secondaire car la diminution prévue de la clientèle scolaire à ces niveaux diminue leur dynamisme de renouvellement. Cet accroissement est plus ou moins rapide selon l'importance et l'âge moyen comparés des différents mouvements qui composent les "apports". Seul le collégial,

(1) Les hypothèses ont été faites après examen des tendances et des caractéristiques connues des mouvements (arrivées, retours, etc.) qui assurent le renouvellement.

TABLEAU 24 : Comparaison des taux bruts de départ obtenus à partir des 3 simulations

SECTEUR PUBLIC

(repérage en fin d'année)

SEXE	NIVEAU	SIMULATION	1971/72	1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
MASCULIN	ELEMENTAIRE	1	.162	.158	.159	.159	.155	.155	.155
		2	.150	.143	.143	.143	.136	.136	.136
		3	.121	.118	.117	.117	.114	.114	.114
	SECONDAIRE	1	.158	.159	.155	.152	.152	.150	.149
		2	.153	.157	.155	.152	.152	.150	.145
		3	.118	.117	.117	.115	.115	.114	.114
	SUPERIEUR	1	.312	.312	.314	.314	.312	.312	.314
		2	.313	.313	.315	.315	.311	.313	.316
		3	.180	.180	.183	.183	.180	.180	.183
FEMININ	ELEMENTAIRE	1	.139	.139	.136	.135	.132	.132	.132
		2	.126	.125	.121	.121	.117	.117	.117
		3	.117	.117	.112	.113	.107	.107	.107
	SECONDAIRE	1	.119	.117	.115	.113	.111	.110	.108
		2	.112	.110	.107	.105	.103	.102	.100
		3	.095	.093	.090	.089	.087	.086	.084
	COLLEGIAL	1	.206	.201	.201	.197	.193	.189	.180
		2	.183	.180	.180	.177	.174	.171	.165
		3	.156	.153	.153	.148	.144	.141	.133
	UNIVERSITAIRE	1	.378	.381	.376	.376	.380	.385	.385
		2	.369	.370	.366	.366	.370	.374	.375
		3	.262	.263	.261	.261	.263	.264	.264

TABLEAU 25: Comparaison des taux bruts de départ obtenus à partir des 2 simulations

SECTEUR PRIVE

(Repérage en fin d'année)

AGE	NIVEAU	SIMULATION	1971/72	1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
MASCULIN	ELEMENTAIRE	1	.388	.387	.387	.384	.390	.388	.387
		2	.415	.414	.412	.408	.421	.420	.417
		3	.402	.401	.398	.391	.417	.415	.410
	SECONDAIRE	1	.298	.296	.295	.294	.293	.291	.291
		2	.283	.280	.280	.281	.278	.275	.275
		3	.246	.244	.243	.242	.241	.239	.240
	COMMERCIAL	1	.337	.330	.329	.329	.328	.329	.329
		2	.314	.303	.302	.302	.301	.302	.302
		3	.259	.236	.232	.235	.231	.233	.223
FEMININ	NIVERSELLE	1	.466	.466	.462	.462	.460	.455	.455
		2	.444	.445	.440	.441	.438	.433	.433
		3	.417	.417	.413	.413	.410	.406	.405
	ELEMENTAIRE	1	.290	.291	.281	.291	.291	.287	.282
		2	.275	.276	.265	.276	.276	.272	.266
		3	.234	.235	.221	.236	.237	.231	.223
	SECONDAIRE	1	.334	.334	.333	.333	.331	.331	.331
		2	.319	.320	.318	.318	.316	.316	.316
		3	.249	.249	.248	.247	.244	.242	.241
	COMMERCIAL	1	.456	.455	.457	.454	.455	.455	.455
		2	.470	.469	.471	.467	.468	.468	.469
		3	.450	.449	.450	.443	.443	.444	.444

pour des raisons inverses, est gratifié d'une baisse marquée de l'âge moyen. Pour le secteur privé, où l'âge moyen est plus élevé, nous avons tenu compte du fait que ce secteur, volontairement ou par nécessité recrute des enseignants assez âgés: l'âge moyen est supposé n'y suivre aucune tendance marquée ni à la hausse, ni à la baisse.

- Au cours des années observées, la variance de l'âge a beaucoup baissé dans les catégories où elle était forte: c'est-à-dire que le personnel, par vieillissement des "apports" et perte des enseignants âgés, tend à être plus concentré autour de l'âge moyen. Cette diminution de la variance est supposée se stabiliser en fin de période pour les niveaux où elle atteint une valeur assez faible: ce qui est surtout valable pour le secteur public. Au privé, et surtout pour le sexe féminin, où la variance était à un niveau très élevé, la diminution de la variance est continue. Ceci provient de l'élimination progressive des enseignants très âgées. Un autre facteur qui plaide en faveur de ces diminutions, est la réduction de l'écart entre l'âge moyen des enseignants "stables" et celui des "apports" (et même une ré-

duction de l'écart entre les différents mouvements qui composent les apports).

- L'âge modal (compris entre 22 et 28 ans) se déplace dans le même sens que l'âge moyen et tend donc à augmenter.

- L'âge de début de l'intervalle d'âges dépend des valeurs et de l'évolution observées (ainsi que de la de la fonction de Pearson utilisée): il varie entre 18 et 22 ans, et augmente avec le niveau;

- L'âge final de l'intervalle d'âges, varie entre 62 et 69 ans et prend ses valeurs les plus fortes au secteur privé.

3. Les résultats

Ces résultats sont présentés dans les tableaux 24, 25 (comparaison des taux bruts), 26,27, et 28 (nombres absolus de départs).

Il faut rappeler encore une fois qu'il s'agit seulement d'une première simulation, obtenue à partir d'une base de données relativement étroite: ces résultats devraient donc, quant à leur fiabilité, connaître une certaine amélioration dans l'avenir.

TABLEAU 26: Comparaison du nombre de départs obtenu à partir des 3 simulations

SECTEUR PUBLIC

(Repérage en fin d'année)

SEXE	NIVEAU	SIMULATION	1971/72	1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
MASCULIN	ELEMENTAIRE	1	620	548	522	504	471	452	436
		2	575	495	470	453	413	397	382
		3	463	407	385	370	346	332	319
	SECONDAIRE	1	3,386	3,640	3,473	3,188	3,038	2,778	2,489
		2	3,372	3,604	3,463	3,198	3,031	2,784	2,484
		3	2,515	2,691	2,603	2,422	2,296	2,12.	1,901
	COLLEGIAL	1	931	1,142	1,426	1,549	1,561	1,621	1,792
		2	935	1,144	1,431	1,553	1,558	1,624	1,804
		3	537	658	830	901	900	935	1,043
FEMININ	MATERNELLE	1	381	357	337	328	315	304	264
		2	344	323	301	294	278	269	233
		3	319	303	279	273	255	247	214
	ELEMENTAIRE	1	3,698	3,288	3,060	2,900	2,726	2,606	2,460
		2	3,463	3,074	2,842	2,696	2,534	2,422	2,282
		3	2,932	2,600	2,396	2,271	2,130	2,036	1,916
	SECONDAIRE	1	3,247	3,348	3,257	3,000	2,780	2,541	2,173
		2	2,890	2,994	2,913	2,693	2,508	2,304	1,997
		3	2,465	2,536	2,467	2,258	2,082	1,900	1,611
	COLLEGIAL	1	484	597	732	794	816	856	942
		2	471	581	712	773	793	833	916
		3	335	412	506	552	564	588	646

TABLEAU 2.1 : Evolution du parc de véhicules utilisés à partir des 3 enquêtes

SECTEUR PRIVE

(Repérage en fin d'année)

			1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
ENSEMBLE	VOITURES PARTICULIERES	1	64	64	58	52	48	44	42
		2	68	69	61	55	52	48	45
		3	66	67	59	53	51	47	44
	VEHICULES UTILITAIRES	1	335	298	336	377	411	450	491
		2	403	318	357	392	418	426	426
		3	365	277	311	339	362	370	371
	DEUX ROUES	1	102	109	125	130	135	148	168
		2	95	100	115	119	124	136	154
		3	78	78	88	92	95	105	119
ENTREPRISES	VOITURES PARTICULIERES	1	47	47	41	41	39	37	32
		2	44	45	39	39	37	36	31
		3	38	42	36	36	35	33	29
	VEHICULES UTILITAIRES	1	218	220	191	179	164	150	139
		2	223	209	180	170	155	142	131
		3	169	173	151	146	133	120	110
	DEUX ROUES	1	374	465	520	568	609	627	627
		2	359	442	497	548	580	597	598
		3	291	316	387	421	448	459	457
	TOTAL	1	138	151	174	179	187	205	233
		2	144	156	179	184	192	211	239
		3	136	149	171	174	182	200	227

TABLEAU 28 : Comparaison du nombre de départs obtenu à partir des 3 simulations

SECTEUR PUBLIC & PRIVE

(Repérage en fin d'année)

SEXE	NIVEAU	SIMULATION	1971/72	1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
MASCULIN	ELEMENTAIRE	1	684	612	580	556	519	496	478
		2	643	564	531	508	465	445	421
		3	529	474	444	423	397	379	365
	SECONDAIRE	1	3,621	3,933	3,809	3,565	3,449	3,228	2,940
		2	3,655	3,922	3,820	3,590	3,449	3,210	2,909
		3	2,760	2,968	2,914	2,761	2,658	2,424	2,272
	COLLEGIAL	1	1,033	1,251	1,551	1,679	1,696	1,769	1,960
		2	1,030	1,244	1,546	1,672	1,682	1,760	1,958
		3	615	736	918	993	995	1,040	1,162
FEMININ	MATERNELLE	1	428	401	378	369	354	341	296
		2	388	368	340	333	315	305	264
		3	361	345	315	309	290	280	242
	ELEMENTAIRE	1	3,933	3,508	3,539	3,226	2,890	3,756	2,599
		2	3,686	3,283	3,022	2,866	2,689	2,564	2,413
		3	3,121	2,778	2,547	2,417	2,263	2,156	2,026
	SECONDAIRE	1	3,603	3,813	3,777	3,578	3,389	3,168	2,800
		2	3,249	3,438	3,400	3,235	3,088	2,901	2,595
		3	2,746	2,882	2,854	2,679	2,530	2,359	2,068
	COLLEGIAL	1	622	748	906	973	1,003	1,061	1,175
		2	613	737	891	957	985	1,044	1,155
		3	473	561	679	726	746	788	873

A partir de ces tableaux, on peut faire les remarques suivantes:

- Le nombre des départs restera encore important parmi le personnel enseignant: ce qui est logique, puisque les probabilités de départ ont été établies à partir d'années où justement, les départs étaient nombreux. Toutefois ce nombre absolu est en baisse puisque l'effectif du personnel est généralement sur le déclin.

- A l'intérieur de chacune des simulations, la modification de la structure par âge tend à favoriser, légèrement, une diminution des taux brut de départ.

- En fait, le choix de la période de régression apparaît, comparé avec l'effet de la structure par âge, comme le facteur décisif de la simulation: c'est lui qui détermine le niveau du taux brut de départ. Celui-ci variant peu par la suite.

- La simulation no 1 (toutes les années sont prises en compte) et la simulation no 2 (on exclut 1968-69/ 1969-70) donnent des résultats comparables.

- La simulation no3 (la dernière période, seule, sert de base à la régression) donne des résultats beaucoup plus faibles. Cela est surtout vrai au secteur public, et tout particulièrement au collégial; mais cela est également valable pour le secondaire. Tout le problème consiste donc à savoir si la période 1970-71/1971-72 est un accident (à la baisse), ou bien si elle marque le début d'une nouvelle tendance du comportement des enseignants. La réponse à cette question ne pourra être donnée que lorsque la comparaison des années 1971-72 et 1972-73 pourra être faite. Quelle que soit cette réponse, d'ailleurs, le modèle devra en tenir compte: il faudra recalculer toutes les équations.

- Les écarts de taux brut, entre les différentes simulations, n'entrainent des différences significatives dans le nombre absolu des départs, que si l'effectif du personnel est important. Pour tous les autres cas, la "fourchette" de prévision est suffisamment restreinte pour servir telle quelle; une seule exception: le collégial.

APPENDICE STATISTIQUE SOMMAIRE

Cet appendice ne sert qu'à rappeler certaines notions de base. Le lecteur voulant approfondir le sujet pourra consulter les ouvrages suivants:

En langue française:

L.T. DAYHAW: "Manuel de statistique" Editions de l'université d'Ottawa- 1969.

G. CALOT: "Cours de statistique descriptive" Dunod, 1965.

G. CALOT: "Cours de calcul des probabilités" Dunod, 1967. (traitement assez mathématiques)

E. MALINVAUD: "Méthodes statistiques de l'économétrie" 2e éd., Dunod, 1971.

J. DUBOS: "Liaisons stochastiques en économie" Dunod, 1971 (avec des exemples d'application).

En langue anglaise:

J. JOHNSTON: "Econometric methods" Mc Graw-Hill, 2e ed, 1972.

Une distribution statistique (nombres absolus, fréquences ou probabilités) se caractérise surtout par son mode (valeur du maximum) et ses moments:

- Moment d'ordre 1 ou moyenne (1) $\bar{x} = \frac{\Sigma\ x_i}{N}$
- Moment centré d'ordre 2 ou variance: $M_2 = \frac{\Sigma\ (x_i - \bar{x})^2}{n}$ dont la racine carrée fournit l'écart-type.
  Ces deux valeurs sont, chacune à leur échelle, d'autant plus faibles que la distribution est peu dispersée autour de la moyenne.
- Les moments centrés d'ordre 3: $\left(M_3 = \frac{\Sigma\ (x_i - \bar{x})^3}{n}\right)$ et d'ordre 4: $\left(M_4 = \frac{\Sigma\ (x_i - \bar{x})^4}{n}\right)$ servent à caractériser la forme de la distribution par l'intermédiaire de:
- l'indice d'assymétrie ("skewness"): $\frac{(M_3)^2}{(M_2)^3}$

  (l'indice est nul quand la symétrie est parfaite)
- l'indice de Kurtose (aspect aigu ou non de la courbe) (2):

  $\frac{M_4}{(M_2)^2}$ qui est calculé par rapport à une courbe de loi Normale ("Mesokurtique") où il est égal à 3. Si la courbe est aiguë ("leptokurtique"), l'indice est supérieur à 3; si la courbe est plus aplatie que celle de la loi Normal ("platykurtique"), l'indice est inférieur à 3.

---

(1) On parle d'"espérance mathématique" pour une variable aléatoire.

(2) Appelé aussi indice d'applatissement.

La régression consiste à estimer les coefficients d'une équation reliant plusieurs variables:

$$Y=aX+bZ+c \longrightarrow \hat{Y}=\hat{a}X+\hat{b}Z+\hat{c}+\varepsilon$$

- les coefficients ($\hat{a}$,$\hat{b}$ et $\hat{c}$) sont appelés "coefficients de régression"
- Y est la variable "dépendante" et X et Z sont les variables "explicatives".
- ε caractérise le résidu aléatoire dont la loi de probabilité est censée avoir une moyenne nulle.
- On peut comparer les variables:
  - 2 à 2, par le <u>coefficient de corrélation simple</u>

$$r=\frac{\frac{1}{n}\Sigma (x_i-\bar{x})(y_i-\bar{y})}{\text{Ecart-type}(x) \times \text{Ecart-type}(y)} \qquad 0 \leq |r| \leq 1$$

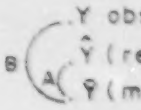
  - 2 à 2, en controlant les autres variables, par le <u>coefficient de corrélation partiel.</u>
  - la variable dépendante par rapport à <u>toutes</u> les variables explicatives ensemble, par le <u>coefficient de corrélation multiple</u>.

- la qualité de la régression se juge par certains indices statistiques:

→ <u>l'écart-type d'estimation</u> ("Standart error of estimate")

$$\sigma_{yx} = \left(\frac{\Sigma (\hat{Y}_i - Y)^2}{n} \right)^{\frac{1}{2}}$$

→ <u>le coefficient de détermination</u> (R^2), qui mesure le gain de précision obtenu par la régression par rapport à l'utilisation de la moyenne conditionnelle.



$R^2 = \Sigma \left(\frac{A^2}{B^2} \right)$; égal à 1 si $\hat{Y} = Y$

→ <u>le F de l'analyse de variance</u>, qui doit être supérieur à une valeur critique F_α, qui est tirée des tables.

▷ Une équation de régression linéaire exige certaines conditions:

- des variables explicatives ne doivent pas être liées entre elles par une relation linéaire (absence de colinéarité).

- les résidus ε doivent être distribués suivant une loi Normale, d'espérance mathématique nulle, et de variance constante (homoscédascité); de plus ils ne doivent pas être autocorrélés.

Ces différentes conditions peuvent être vérifiées par des tests.

ANNEXE

Nous avons regroupé dans cette annexe les tableaux donnant les valeurs (réelles jusqu'en 1971-72, projetées après) de certaines variables (l'âge moyen, l'âge modal, la variance) utilisées dans les simulations.

Tableau no 1 Série des âges moyens utilisée pour la simulation
(en années d'âge)

Catégorie		Année	1967/68	1968/69	1969/70	1970/71	1971/72	1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
SECTEUR PUBLIC	Masculin	Elémentaire	29.9	29.9	29.8	29.8	30.1	30.4	30.4	30.6	30.6	30.8	30.8
		Secondaire	31.5	31.4	31.5	31.9	32.4	32.6	32.8	33.0	33.3	33.6	34.0
		Collégial	32.3	33.3	33.5	32.7	32.4	32.0	31.6	31.2	31.0	30.8	30.8
	Féminin	Maternelle	27.8	27.6	27.7	28.3	29.0	29.4	29.8	30.1	30.3	30.5	30.8
		Elémentaire	31.4	31.7	32.2	32.8	33.2	33.4	33.8	34.0	34.3	34.6	35.0
		Secondaire	32.3	32.1	32.3	32.5	33.0	33.0	33.2	33.5	33.8	34.2	34.8
		Collégial	35.0	33.5	34.5	32.7	32.1	31.7	31.3	31.0	30.8	30.5	30.0
SECTEUR PRIVE	Masculin	Elémentaire	34.5	34.1	32.3	31.8	32.7	32.5	32.8	33.0	33.2	33.4	33.4
		Secondaire	34.3	36.4	36.7	36.5	36.6	36.5	36.2	36.5	36.4	36.3	36.5
		Collégial	35.3	35.6	35.2	36.7	37.0	37.2	37.4	37.6	37.8	38.0	38.0
	Féminin	Maternelle	35.7	35.9	35.6	34.0	34.8	34.8	34.8	35.0	35.3	35.5	35.8
		Elémentaire	35.9	35.6	35.1	34.5	35.1	34.6	34.4	34.1	34.0	34.2	34.5
		Secondaire	38.1	38.4	37.9	37.6	37.5	36.7	36.4	36.2	36.0	35.8	36.0
		Collégial	36.2	36.2	32.7	34.2	38.3	38.0	37.8	37.5	37.1	36.6	35.8

Tableau no 2 Série des variances (de la distribution par âge) utilisée pour la simulation

Catégorie			Année 1967/68	1968/69	1969/70	1970/71	1971/72	1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
SECTEUR PUBLIC	Masculin	Elémentaire	92.9	87.8	83.4	79.9	77	75	72	70	68	68	68
		Secondaire	80.4	77.9	78.2	77.4	76	75	74	73	72	71	70
		Collégial	83.8	91.3	88.0	73.5	73	73	70	68	66	68	70
	Féminin	Maternelle	83.3	72.1	68.7	67.1	65	63	61	60	59	58	60
		Elémentaire	133.0	126.3	122.8	118.1	113	108	104	100	96	95	94
		Secondaire	116.9	110.7	109.9	106.1	104	100	98	95	93	90	90
		Collégial	135.6	118.6	110.6	98.1	95	90	85	80	78	75	75
SECTEUR PRIVE	Masculin	Elémentaire	162.2	149.0	126.2	107.8	100	95	90	86	83	80	78
		Secondaire	121.0	145.4	138.2	139.8	135	132	129	126	125	123	125
		Collégial	98.4	97.4	97.9	98.5	96	94	92	90	88	86	85
	Féminin	Maternelle	140.4	140.9	148.2	146.9	145	142	138	132	125	120	115
		Elémentaire	151.1	155.2	145.2	134.2	128	120	112	108	105	102	100
		Secondaire	166.5	172.7	170.4	166.4	162	158	155	150	142	135	130
		Collégial	121.5	126.6	94.8	100.8	102	105	100	95	92	89	86

Tableau no 3 Série des modes des distributions par âge utilisée pour la simulation

(en années d'âge)

Catégorie			Année 1967/68	1968/69	1969/70	1970/71	1971/72	1972/73	1973/74	1974/75	1975/76	1976/77	1977/78
SECTEUR PUBLIC	Masculin	Elémentaire	24	25	24	24	24	25	25	25	26	26	26
		Secondaire	25	25	26	25	25	26	26	26	27	27	28
		Collégial	26	26	27	27	27	27	26	26	27	27	26
	Féminin	Maternelle	21	22	23	23	24	24	25	25	26	26	26
		Elémentaire	23	23	24	25	25	26	26	27	27	28	28
		Secondaire	24	23	24	24	24	25	25	25	25	26	26
		Collégial	25	24	26	25	25	25	24	24	24	25	25
SECTEUR PRIVE	Masculin	Elémentaire	25	24	27	24	24	24	25	24	23	23	24
		Secondaire	26	27	26	25	26	27	27	26	27	28	28
		Collégial	26	27	27	28	27	28	28	29	29	30	30
	Féminin	Maternelle	24	27	27	24	24	25	25	26	27	27	28
		Elémentaire	25	25	26	24	25	25	24	25	25	26	26
		Secondaire	24	25	26	25	25	25	26	26	27	27	27
		Collégial	25	26	26	27	27	27	27	28	28	28	28